United Nations

A/70/686

General Assembly

Distr.: General
21 March 2016

Original: English

Seventieth session
Agenda item 140
Joint Inspection Unit

Analysis of the evaluation function in the United Nations system

Note by the Secretary-General

The Secretary-General has the honour to transmit to the members of the General Assembly the report of the Joint Inspection Unit entitled "Analysis of the evaluation function in the United Nations system".

Please recycle

JIU/REP/2014/6

ANALYSIS OF THE EVALUATION FUNCTION IN THE UNITED NATIONS SYSTEM

Prepared by

Sukai Prom-Jackson
George A. Bartsiotas

Joint Inspection Unit

Geneva 2014

EXECUTIVE SUMMARY

Analysis of the evaluation function in the United Nations system

JIU/REP/2014/6

In the current context of resource constraints and national and global challenges that require new approaches in development and humanitarian assistance, United Nations system organizations are increasingly faced with the need to account for the use of resources and demonstrate results and the added value of their work. They are operating in an environment that calls for impartial, valid and credible evidence on the relevance, added value, effectiveness, efficiency, impact and sustainability of their policies, programmes and activities. In such an environment the evaluation function provides a platform for organizations to respond to these demands across the system.

Evaluation is one of the main instruments that support the United Nations system in addressing accountability for results and added value, for learning and knowledge development, strengthening its leadership role in global governance, and instituting reforms that influence the lives of people worldwide. Thus the continuous development of the evaluation function is critical to the United Nations system's ability to achieve its objectives, to account for success, and bring about necessary changes to improve international development and governance.

The present report contains the results of a Joint Inspection Unit (JIU) system-wide study of the evolution, development and advancement of the evaluation function in the United Nations system conducted in 2013. The study addressed the following questions: Has there been progression in the growth and development of the evaluation function in the United Nations system? What is the level of development of the evaluation function? What is its capacity to support United Nations system organizations and its response to demands, changes and challenges? How has it added value? What alternative approaches exist for an effective evaluation function that serves the United Nations system in the current context?

The study sought to contribute to on-going efforts across the system, directed at strengthening the capacity of the evaluation function to meet professional standards, address emerging challenges and play a role in enhancing the value of the United Nations system. The focus is on the central evaluation function, which generally supports overall corporate-level policy and strategic decision-making. The study also provides a rapid review of the decentralized evaluation function, which operates outside the central evaluation unit and is generally embedded in programme and operational units throughout the United Nations system supporting line management decision-making.

The study covered the evaluation function of 28 United Nations system organizations involved in development, humanitarian and normative work worldwide: that of 12 funds or programmes, 12 specialized agencies, 3 other United Nations entities and the central evaluation function of the United Nations Secretariat. It did not, however, include the embedded evaluation functions of the United Nations Secretariat departments or the regional commissions as they had already been addressed in a recent evaluation study conducted by the United Nations Secretariat itself.

The report provides nine recommendations: seven to executive heads of United Nations system organizations and two to United Nations system legislative bodies.

The following is a summary of the main conclusions and the supporting findings and key issues addressed in the report.

A. Central evaluation function

1. ***The central evaluation function has grown through the years, striving for quality and efficiency, but the level of commitment to evaluation across the United Nations system is not commensurate with the growing demand for and importance of the function.***

Assessed over a 36-year time span (1977–2013), the central evaluation function of United Nations system organizations has changed with regard to roles, structure, systems and standards. It has moved from a predominant role of oversight over and quality assurance of decentralized evaluations to focusing on supporting broad and strategic corporate-level decision-making. The emerging global challenges faced by the United Nations system call for greater involvement of the central evaluation function in cross-cutting and multi-sectoral types of evaluation that provide more integrated and holistic solutions for the system.

Organizations have remained pragmatic and cost-conscious in developing their evaluation functions. That applies both to the positioning of the function in the organizational structure and to the business models developed to enhance the efficiency and quality of evaluation reports. It is evident, however, that the level of resources allocated to the central evaluation function has not changed significantly over the years in ways that are commensurate with demand.

The initial creation and development of the central evaluation function was influenced by several JIU reports on organizational management, administration and governance. General Assembly resolutions have highlighted the importance of the evaluation function and the expanded role it could play as an instrument of accountability and learning. The United Nations Evaluation Group (UNEG) has also played a significant and highly visible role over the past eight years in enhancing the professional development of the function and the harmonization of evaluation methodologies, thus providing an effective platform for the advancement of the function.

2. ***The quality of evaluation systems, mechanisms, processes and outputs varies across the United Nations system. Organizations can be categorized into four clusters according to the level of development of their respective evaluation functions. In most cases, the level of development is affected by the size of the organization, the resources allocated to evaluation, and the structural location of the function.***

This JIU system-wide study assessed the status of development of evaluation in 28 organizations by analysing five core components that define an effective evaluation function: (a) the adequacy of the enabling environment; (b) independence, with inclusion and impartiality to enhance credibility; (c) quality, to enhance credibility; (d) the utility of the function focused on the use of evaluation findings and recommendations; and (e) relevance and adaptability to address organizational demands, and readiness to respond to change and challenges.

Based on the analysis of the level of maturity of their central evaluation function, organizations are categorized into four clusters: nine organizations are considered high performers, six organizations average, and nine below average, while four are rudimentary and have no defined evaluation function. These levels of maturity were determined based on a systematic and standardized assessment using a benchmarking framework (i.e., the JIU maturity matrix for the central evaluation function). The framework has 66 performance indicators benchmarked against a combination of inputs including UNEG norms and standards, JIU parameters from previous reports, and inputs from other international development actors assessing organizational effectiveness.

Most large and medium-sized organizations are assessed as having a higher level of development of their evaluation functions than small organizations. Organizations in which the function is stand-alone are ranked higher than organizations in which the central function is co-located, meaning that the evaluation function is either within the oversight offices concerned with evaluation, audit, inspection and investigation, or within executive management offices concerned with policy, strategic planning and research. Organizations with a predominantly normative mandate perform less well than organizations with a predominantly developmental focus. There are exceptions to the above suggesting that an agency can overcome organizational obstacles when there is the will and determination to develop a high-quality evaluation function.

There are four organizations that do not have a central evaluation function for reasons that have not been fully justified. Two of those organizations have initiated the process of developing an evaluation policy or have begun consultations to develop a fully-fledged central evaluation function.

On the basis of analysis of the five core components of the evaluation function, the following conclusions have been drawn from across the 24 United Nations system organizations that have central evaluation functions. The variations that exist among organizations must be borne in mind when considering the cross-cutting conclusions.

Enabling environment

3. ***In developing their evaluation functions, organizations have focused on responding to demands for accountability and have not fully addressed other important elements, such as developing a culture of evaluation and using evaluation as a learning instrument for the organization. That limits the sustainability of the function and the added value of evaluation. Absent from the institutional framework is an overarching vision and strategy for evaluation that is anchored in the organization's mandate, corporate goals and reform agendas, thereby resulting in a balanced approach to the development of the function.***

Member States and legislative bodies have played a significant role in calling for the establishment of evaluation functions, ensuring compliance with policy implementation, and requiring policy updates. The main focus of Member States has been on accountability for resources and results. Organizations have been responsive to those demands, but they have not placed equal emphasis on the development of the learning function and a culture of evaluation. They have not made evaluation an integral part of the fabric of the organization or acknowledged its strategic role in going beyond results or performance reporting. An imbalance between the accountability and learning goals of evaluation does not enhance the added value and sustainability of the evaluation function, particularly in the current fast-changing global environment in which the flow of knowledge is moving rapidly across boundaries, requiring dedicated and nimble knowledge workers with critical and divergent thinking capacities.

Most organizations have evaluation policies in place. The policies reveal a high level of adoption of the norms and standards promulgated by UNEG. A large number of those policies, however, reflect a mechanical adoption of the UNEG norms and standards, without their being adapted to the institutional context. Similarly, most organizations do not have an overall vision and strategy for the central evaluation function that address strategic alignments such as with other oversight and learning systems, with the decentralized evaluation function, with other United Nations organizations or with national systems. The absence of a broad strategic vision limits adaptability and decision-making and presents risks associated with ad hoc approaches in the development of the function.

4. ***The United Nations system evaluation function is under-resourced and overstretched.***

The staff ratio of 0.2 per cent, defined in terms of the percentage of professional evaluation staff to overall organizational staff, and the funding level of the evaluation function at about 0.3 per cent of organizational expenditure are assessed to be inadequate to address the growing demand for coverage, quality and institutional support. This limited capacity leaves little room to respond to the vast and growing demands on the function, or to enhance its broader role in supporting the United Nations system as whole.

There is a need in most organizations to increase the resourcing of the evaluation function and to develop appropriate costing norms and frameworks to guide budget allocation for evaluation. There is also a need in some organizations for more efficient use of staff to focus on carrying out the core evaluation function, as opposed to other oversight or management functions. Also significant in enhancing the efficiency of the central function is the development of institutional support functions such as better systems of monitoring and data quality and well-designed programmes with clear objectives, well established results framework and performance indicators all of which facilitate the evaluability of the programmes.

Of equal importance to enhance both efficiency and added value is exploring alternatives in programme delivery to include greater interdependence and joint work in evaluation systems development and an increase in joint evaluations. The on-going debate on the post-2015 development agenda indicates that the United Nations system can best survive and flourish when it breaks down so-called “silos”. This implies that the evaluation function of the United Nations system should restructure itself by reversing the predominant approach of trying to “go it alone” in each organization, particularly small organizations, and should engage a wider range of partners and institutions. Such an evaluation function would play a greater role in providing multi-sectoral and cross-cutting information that would support holistic and strategic decision-making in the United Nations system.

Credibility of the function: independence and quality

5. ***Progress has been made in enhancing the credibility of the function by increasing its independence and issuing better quality evaluation reports, but more needs to be done, in particular to address structural independence, the external quality assurance of the function, and the quality of reports.***

The independence of the function is the component with the highest level of development of the five components that define a high-quality evaluation function. The results of an assessment of five criteria of independence (structural, functional, professional, technical, inclusion and behavioural independence), show that **two** organizations operate with high levels of independence, **16** organizations at above-average levels, **six** at average levels and six at well below an acceptable level of independence. Great progress in enhancing independence has been made through the role played by both evaluation managers and evaluation staff in maintaining high standards of professional and personal integrity with regard to independence. Progress has also been made by senior management in safeguarding independence. Independence with inclusion is an important and much valued characteristic of the United Nations evaluation system. For a multi-cultural, multi-dimensional and inter-governmental body like the United Nations system, it enhances trust and transparency. Most of the evaluation units of United Nations system organizations have done well in enhancing stakeholder involvement in evaluation while also safeguarding independence. Challenges remain in enhancing structural independence and principally in addressing the independence of the head of the evaluation function and the reporting lines. Other system-wide issues to be addressed for enhancing independence, while also enhancing technical quality and credibility of the function, include addressing the regional and gender imbalance among heads of evaluation units, and ensuring the professional evaluation expertise of unit heads.

Regarding the professional and technical quality of the function, the main strengths are in staff competencies and behaviours, although more needs to be done to support the professional development and excellence of staff. Involving stakeholders to strengthen the content, value and meaningfulness of evaluation reports is a common practice in a large number of organizations. Significant challenges include: (a) the limited use of independent, external quality assurance mechanisms for the quality assessment, or meta-evaluation of reports; (b) the lack of solid information about the quality of evaluation reports; and (c) the current low quality of reports. The UNEG peer review mechanism, which was initiated in 2006 and assesses both the function and the quality of reports, has been implemented only in nine organizations. External assessments of the quality of reports have some common elements, but they use different measurement systems. There has been an increase in external assessments of the evaluation function in recent years by bilateral organizations. Alignment is needed among all existing mechanisms for the quality assessment of the evaluation function and its products to avoid duplication, enhance reliability, limit transaction costs and have an impact on harmonization and efficiency.

Utility: use of evaluation and effect of use

6. ***Organizations are not predisposed to a high level of use of evaluation to support evidence-based policy and decision-making for strategic direction setting, programmatic improvement of activities, and innovations.***

The study found that the use of evaluation reports for their intended purposes is consistently low for most organizations. Even organizations in which the evaluation function is considered to perform well manifest, only an average level of use of reports by the intended audience. Low level of use is associated with an accountability-driven focus and the limitations noted above on the role of the function in the development of the learning organizations. There is a need to improve the systems in place for assessing the use of evaluation. Likewise, better systems are needed for assessing the impact on organizational effectiveness in using evaluation. Currently, the systems used to assess that impact are rudimentary and ad hoc. The limited information available on the use and impact of evaluation reports makes it difficult to provide analyses for a full understanding of the overall value of the function. The United Nations system is, however, not alone in having a low level of performance regarding the use of evaluation to influence decisions and turn learning into action. Prevailing evidence shows that other development partners have the same problem, which suggests that a concerted effort is needed to address the problem and enhance understanding of the value of the function.

Given the interest in the use and hence the value of the function, the study examined in an exploratory fashion the overall value for money of the function. Analysis showed that the nine organizations in which the evaluation function is the most developed account for almost 70 per cent of the total organizational expenditure of the 28 JIU participating organizations. While it is indicative only, the evidence suggests that the evaluation function of the United Nations system is good value for money, and could be more so with the continuing development of the function, including enhanced use of evaluation reports.

Readiness to support change

7. ***Factors such as the different mandates and business models of United Nations organizations, the tendency towards the "silo effect" — not thinking in broad strategic terms or considering wider alliances and alignments within and across United Nations agencies — and the generally low level of evaluation capacity in a number of organizations hinder the ability of the evaluation function to play a key role in driving change in the United Nations system. Associated with that is the disaggregated functional structure of the United Nations system.***

The level of readiness of the evaluation function to address change and the emerging challenges of the United Nations system is rudimentary in the majority of cases. There is harmonization and coherence in methodology and standards based on the wide range of normative products of UNEG. That has not, however, translated into extensive joint work by the different organizations. There has been an increase in joint evaluations, mainly regarding gender, food security and in humanitarian areas, but they are limited by systemic constraints. The evaluation function of most organizations has yet to come fully to terms with changing global dynamics, the demands for interdependence and inter-connectivity in real terms and how they could be used as an opportunity for restructuring the function. Many organizations are still looking for best practices in traditional ways of doing things when what would perhaps best serve the United Nations system is a reflection on alternative integrative and collaborative systems focused on connections among organizations and sub-systems to optimize performance.

The topics of inclusion (addressing diversity, gender equality and human rights), and national ownership and capacity for evaluation are two important challenges in the area of development. Enhancing gender equality in the management, design, conduct and reporting of the findings of evaluation has been driven by the persistent engagement of UN-Women. Nine of the organizations reviewed manifest a high level of performance in mainstreaming gender in evaluation.

The development of national evaluation capacity (highlighted in several resolutions) is generally not perceived by many organizations as a cross-cutting mandate. Besides its value for the governance of countries, national evaluation capacity development is recognized by most United Nations system organizations for its effect on the utility and sustainability of their evaluation function, no matter what their organizational size, structure, activities or field presence. Support of the evaluation function for national evaluation capacity development is, however, rudimentary. Organizations manifest varying forms of engagement with national institutions and experts in evaluation to enhance either capacity for evaluation or ownership of evaluation results. Only three organizations implement activities to support national capacities for evaluation based on a strategic approach.

Even though it has been argued that strengthening national capacities for evaluation is a programmatic responsibility, demands for support from the central evaluation function which houses evaluation experts has not abated. These, and several other increasing demands on the central evaluation function raise questions about the capacity and ability of the United Nations evaluation system to deliver as expected. The dynamism and commitment of heads and staff of evaluation units have been a major driving force in improving the quality and value added of evaluation, but are not, on their own, adequate to ensure the sustainability and relevance of the function. Heads of evaluation units have now embarked on the implementation of a new UNEG strategy to address some of the challenges, but full reflection on alternatives, non-traditional approaches, and support is needed from all stakeholders, including management, Member States and private-sector partners.

B. Decentralized evaluation function

8. *The absence of an overarching and well-defined institutional framework, based either on evaluation policies or results-based management, makes the decentralized function tenuous.*

Decentralized evaluations are planned, managed and conducted outside the central evaluation unit. They are embedded within programme and management units responsible for the planning and implementation of projects and programmes. They are structurally not independent of line management. That applies to evaluations planned and managed at the discretion of project managers as well as those commissioned or contracted out to external consultants. Today, most organizations do not have a robust institutional framework to support the decentralized function. Only 11 organizations are assessed as having defined systems or elements of a system for decentralized evaluations.

The study analysed the various archetypes and models of decentralized evaluation in the United Nations system and identified two main models of decentralized evaluation systems.

The first model is the ad hoc system without a defined institutional framework which continues the tradition of discretionary self-evaluation of projects conducted by project staff at headquarters and in field offices. That is the practice in all but one of the 28 organization and involves a vast number of project managers who, serving as a knowledge force, are producing a substantial number of project evaluations. The quantity and quality of their evaluation work is unknown, therefore reducing the strategic utility of such evaluations.

The lack of attention paid to the ad hoc system of decentralized evaluations suggests that the United Nation system is missing opportunities to make better use of the assets from the evaluation reports produced and to engage staff involved in such evaluations that are conducted all over the system. This is an important issue in the current context, given the pivotal role of such staff in responding to the need for providing timely evidence on the ground and making immediate changes where it matters the most. The significance of this issue has been highlighted in the debates on the post-2015 development agenda where the response to fast change and complex challenges require continuous and formative evaluation for correction in the course of implementation.

The second model consists of decentralized evaluation systems with a defined plan, a quality assurance system and systematic reporting. As mentioned above, only 11 of the 24 JIU participating organizations were assessed to have such systems and they do so to varying degrees and at different levels of institutionalization. For 10 of the organizations, the decentralized function is assessed to currently operate out of 923 points (i.e. either evaluation units or evaluation tasks assigned to monitoring and evaluation specialists or focal points) within programme and technical departments at headquarters, and at regional and over 150 country offices producing over 640 reports per year. On average, 64 per cent of the reports (of this group of 10 organizations) are output-level evaluations and 10 per cent are outcome evaluations. There are 12 impact evaluations, while the rest of the evaluations are of unknown type.

The study collected further evidence for those 10 organizations and the observations below apply to those 10.

9. ***Policies, norms and standards, and resources for decentralized evaluations are inadequate.***

The decentralized function for the sample of 10 organizations reviewed is missing an appropriate normative framework that is "fit for purpose" for the role and comparative and added value of the decentralized evaluation. The UNEG framework for norms and standards does not completely provide for the needs of the decentralized function. However, professional networks are being established to address decentralized evaluation, its value systems and the rules of the game, its key issues and challenges and the role it should play for the benefit of individual organizations as well as for the system as a whole. Such networks have the potential to greatly enhance the development of the function and they need to be expanded, empowered and supported. A concerted United Nations effort across agencies is needed to enhance the development of such networks.

Existing organizational evaluation policies have guided to some degree the development of the decentralized function. These policies have not however addressed two important areas: (i) coherence, and (ii) alignment with national evaluation. With regard to their level of development, these areas are among the weakest of the decentralized function.

Funding for decentralized evaluation has come from project and programme budgets in some organizations. That offers some stability as opposed to situations when there is dependence on extra-budgetary financing. However, the amounts provided for conducting evaluations are significantly low relative to the financing of decentralized evaluation in other international development entities.

10. ***The credibility of the decentralized function has improved, but more remains to be done.***

A significant number of measures have been put in place to limit bias and enhance the impartiality of evaluation and this has enhanced credibility. The quality of decentralized evaluation reports has improved, but more needs to be done to enhance the systems and staffing needed to further improve quality. Also, one would have expected greater coherence at the country level and particularly in the context of the unifying framework offered via the United Nations Development Assistance Framework (UNDAF). This has not been the case in many instances even 10 years after UNDAF was introduced. Coordination or collaboration in decentralized evaluation is limited. Coordination between the decentralized and the central evaluation function is almost non-existent in country-level evaluations.

11. ***Use of and learning from decentralized evaluation is limited by an organizational culture which is focused on accountability and responsiveness to donors.***

Like the central function, the decentralized function is highly responsive to demand for accountability. It is also focused on the formative improvement of projects and programmes, which is considered a prime value of the decentralized function. However, the function has not sought to enhance the empowerment of staff and promote learning from evaluations. Even for the ad hoc evaluations that are being conducted by a vast number of project managers, the indication is that most are demand-driven by donors as part of a protocol agreement tied to extra-budgetary financing.

12. ***The decentralized function has an important role to play in addressing current and emerging changes and United Nations system reforms, but it has to be made an integral part of an overall strategic approach to the development of the evaluation function for it to make a contribution.***

The fast pace of development requires just in time and continuous evidence for directing the process of change and development in achieving desired outcomes and impact. Decentralized evaluation operates as part of project implementation thus plays a significant role in providing recurrent strong evidence for correction or the adoption of alternatives. The decentralized evaluation function of the United Nations system has various models, is not fully supported by a well-defined institutional framework and its quality and added value is unclear. There is a need to develop a strategic focus and plan for the enhanced role and comparative value of decentralized evaluations to support the United Nations system at organizational as well as at system-wide level.

Recommendations

Recommendation 1: The executive heads of United Nations system organizations in which the central evaluation function is co-located and integrated with other oversight functions or integrated with the executive management functions should enhance the function and ensure its quality, integrity, visibility and added value.

Recommendation 2: The executive heads of United Nations system organizations should adopt a balanced approach in addressing the purpose of evaluation for accountability, and for the development of a learning organization that has the appropriate incentive systems for innovation, risk-taking and the use of multidisciplinary perspectives.

Recommendation 3: The legislative bodies should request the executive heads of United Nations system organizations to develop comprehensive budget frameworks

and resource allocation plans for their respective evaluation functions, based on the cost of maintaining an effective and sustainable evaluation function that adds value to the organization. The plans should be submitted for consideration to the legislative bodies within existing budgetary and reporting mechanisms and processes.

Recommendation 4: The legislative bodies should direct the executive heads of United Nations system organizations to review and revise, as necessary, existing policies for the appointment of the heads of evaluation offices, in order to enhance independence, integrity, ethics, credibility and inclusion, with due regard to the following criteria:

- **Term limits should be established for a single non-renewable term of office of between five and seven years, with no possibility for the incumbent of re-entry into the organization;**
- **The head of evaluation should have qualifications and substantial experience in evaluation, complemented by experience in the related fields of strategic planning, basic and operational research and knowledge management, and should have excellent management and leadership attributes.**

Recommendation 5: The Secretary-General of the United Nations, in his capacity as chair of the United Nations System Chief Executives Board for Coordination (CEB), should request UNEG to collaborate in developing a robust and harmonized quality-assurance system for the evaluation function across the United Nations system.

Recommendation 6: The executive heads of United Nations system organizations should make the use of evaluation an organizational priority based on a well-defined vision, strategy and results framework for the evaluation function, and report to their legislative bodies on the level, nature and impact of use of evaluation.

Recommendation 7: The executive heads of United Nations system organizations should request evaluation offices to reassess the policies, strategies and priorities of the evaluation function. They should strategically position the evaluation function in their respective organizations so as to enhance its relevance in enabling the United Nations system to address current changes and challenges, and to achieve impact and sustainability.

Recommendation 8: The Secretary-General, in his capacity as chair of the CEB, should initiate steps and support innovations for collaboration among United Nations system organizations and with other partners in strengthening national capacities for evaluation addressing accountability, learning, and knowledge development of both national and global value.

Recommendation 9: The executive heads of United Nations system organizations should develop the institutional framework and necessary support systems to enhance the quality and added value of decentralized evaluation and the role it could play in supporting the United Nations system to address emerging challenges, including those of the post-2015 development agenda, and to enhance coherence and alignments in evaluation within and across United Nations system organizations, and with national institutions.

CONTENTS

Chapter		Paragraphs	Page
	EXECUTIVE SUMMARY		3
	ABBREVIATIONS		15
I.	**INTRODUCTION**	1-35	16-28
	A. Overview	1-5	16
	B. Background: conceptual framework of the evaluation function — demand and the development of a high quality evaluation function	6-19	17
	C. Methodology	20-34	22
	D. Acknowledgements	35	27
II.	**THE CENTRAL EVALUATION FUNCTION: EVOLUTION, GROWTH AND LEVEL OF DEVELOPMENT**	36-139	29-71
	A. Changes in development: purpose, structural arrangements, coverage and management	36-50	29
	B. Organizational level performance: patterns among and variations between organizations	51-60	35
	C. Components of the function: strengths, challenges and systemic constraints	61-139	41
	Component 1: The enabling environment, institutional framework and support systems	62-85	42
	Component 2: Independence with inclusion for enhanced credibility	86-98	51
	Component 3: Quality– technical and managerial rigor for enhanced credibility	99-105	57
	Component 4: Utility - use and impact of use	106-114	59
	Component 5: Relevance and readiness to support United Nations organizations and system-wide reforms and to address emerging changes and challenges	115-139	63
III.	**THE DECENTRALIZED EVALUATION FUNCTION**	140-189	72-87
	A. Introduction	140	72
	B. Institutional and normative framework	141-152	73
	C. Policies, implementation and institutionalization	153-161	76
	D. Resources	162-164	79
	E. Credibility (impartiality and quality) and utility	165-175	81
	F. Readiness for change and challenges: evolving demands, capacity and readiness	176-189	84

ANNEXES

		Page
I.	Results framework of the evaluation function in the United Nations system	88
II.	Components and sub-components of the evaluation function in the United Nations system	89
III.	JIU participating organizations	90

IV (a). Architecture of the evaluation function in the United Nations system 91
IV (b). Architecture of the evaluation function of the United Nations system (details on structural arrangements for the United Nations Secretariat) 92
V (a). Definition of selected terms of the United Nations system evaluation function: central and decentralized evaluation 94
V (b). Decentralized evaluation function – models and archetypes in the United Nations system 98
VI. Business models used by 23 United Nations system organizations 99
VII. Demands on the evaluation function beyond the conduct and presentation of evaluation reports 100
VIII. Evolution of evaluation offices in JIU participating organizations, 1977–2013 101
IX. The learning organization: culture for results, accountability, double loop learning, sharing and use of evaluation 103
X. Leadership attributes and functions 104
XI. Relationship between level of non-core financing, safeguards for effective and independent use, and level of development of the function 105
XII. Readiness for global changes and challenges including the imperatives of the post-2015 development agenda for sustainable development 106
XIII. Overview of actions to be taken by participating organization on the recommendations of the Joint Inspection Unit 107

ABBREVIATIONS

CEB	United Nations System Chief Executives Board for Coordination
FAO	Food and Agriculture Organization of the United Nations
IAEA	International Atomic Energy Agency
ICAO	International Civil Aviation Organization
ILO	International Labour Organization
IMO	International Maritime Organization
ITC	International Trade Centre
ITU	International Telecommunications Union
JIU	Joint Inspection Unit
OHCHR	Office of the High Commissioner for Human Rights
OIOS	Office of Internal Oversight Services
RBM	Results-based management
UNAIDS	Joint United Nations Programme on HIV/AIDS
UNCTAD	United Nations Conference on Trade and Development
UNDAF	United Nations Development Assistance Framework
UNDP	United Nations Development Programme
UNEG	United Nations Evaluation Group
UNEP	United Nations Environment Programme
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNFPA	United Nations Population Fund
UN-Habitat	United Nations Human Settlements Programme
UNHCR	Office of the United Nations High Commissioner for Refugees
UNICEF	United Nations Children's Fund
UNIDO	United Nations Industrial Development Organization
UNODC	United Nations Office on Drugs and Crime
UNOPS	United Nations Office for Project Services
UNRWA	United Nations Relief and Works Agency for Palestine Refugees in the Near East
UN-Women	United Nations Entity for Gender Equality and the Empowerment of Women
UNWTO	United Nations World Tourism Organization
UPU	Universal Postal Union
WFP	World Food Programme
WHO	World Health Organization
WIPO	World Intellectual Property Organization
WTO	World Trade Organization

I. INTRODUCTION

A. Overview

1. The present report describes and analyses the evolution and development of the evaluation function in the United Nations system. The purpose of the report is to help strengthen the capacity of the evaluation function to meet professional standards and address emerging and global challenges across the system.

2. The report follows up on previous comprehensive JIU reports on the evaluation function of United Nations organizations that were issued between 1977 and 1998.[1] It complements the JIU system-wide studies that were conducted between 2005 and 2011 on oversight and accountability,[2] results-based management (RBM)[3] and knowledge management.[4] Moreover, part of the current JIU business strategy is to further enhance its work and contribution to the development of the evaluation function evaluation. Via this report, the JIU seeks to contribute to three major developments: supporting the implementation of the newly developed 2013 United Nations Evaluation Group (UNEG) Strategy; feeding into plans for the 2015 International Year of Evaluation; and enhancing the readiness, strategic positioning and prioritization of the evaluation function of the United Nations system in the post-2015 development agenda.

3. The present report is complemented by extensive supporting materials that can be found on the JIU website and should be read in conjunction with the report. They include annexes providing detailed summaries of data analyses in tables and graphic illustrations. The data provide substantiating evidence for the conclusions reached in the report and include details of the performance of the various organizations, thus providing a rich information base for benchmarking. The supporting materials also include the assessment tools that were used for data collection, including: the questionnaires that were administered for the central and decentralized systems: questionnaire for organizations without a defined function, and the JIU maturity matrix and methodology used to assess the level of development of the central evaluation function. The maturity matrix was completed by the JIU for each of the 24 United Nations system organizations and was provided to the organizations in the first quarter of 2014.[5] Some organizations have reported using the JIU assessments on the matrix to address areas of

---

[1] Relevant JIU reports issued between 1977 and 1998: JIU/REP/98/2; JIU/REP/88/6; JIU/REP/85/10; JIU/REP/83/5; JIU/REP/82/12; JIU/REP/82/10; JIU/REP/81/6; and JIU/REP/77/1.

[2] JIU reports that complement the present report: The investigations function in the United Nations system (JIU/REP//2011/7); Accountability frameworks in the United Nations system (JIU/REP/2011/5); The audit function in the United Nations system (JIU /REP/2010/5); Knowledge management in the United Nations system (JIU/REP/2007/5); and Oversight lacunae in the United Nations system (JIU/REP/2006/2).

[3] Overview of the series of reports on managing for results in the United Nations system (JIU/REP/2004/5).

[4] JIU/REP/2011/7, JIU/REP/2011/5, JIU/REP/2010/5 and JIU/REP/2006/2.

[5] The JIU maturity matrix of the Central Evaluation Function: An Objective and Standardized Approach for the Assessment of the Level of Development of the Evaluation Function of Organizations in the United Nations System, JIU, 2013.

improvement that were identified, some have used it as part of on-going UNEG peer reviews, and others have used it to update their evaluation policies.

4. In accordance with article 11.2 of the JIU statute, the present report was finalized after consultation among the Inspectors so as to test its conclusions and recommendations against the collective wisdom of the Unit.

5. In order to facilitate the implementation of the recommendations contained in the report and the monitoring thereof, annex XIII contains a table indicating whether the report is submitted to the organizations concerned for action or for information. The table identifies the recommendations that are relevant to each organization, specifying whether they require a decision by the organization's legislative or governing body or can be acted upon by the organization's executive head.[6]

B. Background: conceptual framework of the evaluation function — demand and the development of a high-quality evaluation function

Purpose of evaluation: accountability, improvement and responding to growing demand to support reforms and global changes

6. It is well recognized that the capacity for reform and the continued relevance of the United Nations system hinge to a considerable extent on its ability to achieve its objectives, to account for success and to bring about necessary improvements. They also hinge on enhancing its ability to learn and apply alternatives, to engage rigorously in compiling knowledge and good practices, and to use them to fully understand the national, regional and global environments in order to bring about necessary changes in international development and governance.

7. Evaluation plays a key role in enabling the United Nations system to respond to the demands for accountability for results and for added value from its activities. Evaluation is judgement of the value or worth of the activities of the United Nations system. It asks the questions: **Are we doing the right things** (i,e, are they relevant, responsive, appropriate and innovative and not duplicative)? **Are we doing things right** (i.e. are we doing things in an efficient, coherent and inclusive way)? **Are we achieving results** and making a difference and adding value (i.e. are we effective in results attainment and what is the level of attribution or contribution and significance of our work)? **Are we ensuring that we contribute to impact and that our work is sustainable?** Evaluation thus seeks

---

[6] It is important to note that the recommendations are not intended to be specific, measurable, attainable, relevant and timely (SMART), since that would be tantamount to defining project parameters and hence limiting the capacity for independent assessment in the future. Given the variations among the organizations, the recommendations allow for a professional approach in developing the best course of action for each organization. The report provides substantial information that could guide the development of an appropriate response to each recommendation, tailored to organizational specificities, while respecting the standards, goals and integrity of the evaluation function.

"to determine as systematically and objectively as possible the relevance, efficiency, effectiveness and impact of an activity in the light of its goals, objectives and accomplishments".[7]

8. Evaluation explains not only the "what" but also the "why and how" and "in what context". In so doing, evaluation provides a good basis to support management in taking necessary decisions and actions to improve policies, strategies, projects, programmes and operational and normative activities. The main strength of evaluation lies in the rigour of its methodology, using systematic and critical inquiry to provide evidence that is robust, valid, reliable and credible and can be used with confidence in decision-making.

9. It is the reflective and divergent pattern of thinking of the evaluation profession that makes evaluation an important tool for building a culture that is best described as **the learning organization.** That is the organization that acquires knowledge and innovates fast enough to survive and thrive in a rapidly changing environment that requires transformative changes. Learning organizations: (a) create a culture that encourages and supports continuous employee learning, critical thinking via evaluation, and risk taking with new ideas; (b) allow mistakes and value employee contributions; (c) learn from experience and experiment; and (d) disseminate the new knowledge throughout the organization for incorporation into day-to-day activities. In a learning organization, people are engaged in "generative learning that enhances the capacity to create and they are continually learning to see the whole together".[8]

Growing demands and implications for the evaluation function

10. Over the years, the demands on the evaluation function of the United Nations system have expanded and evaluation has been placed at centre stage in the debate on the relevance of the United Nations system. Evaluation is seen as necessary to support the response of the United Nations to increasing demands not only for accountability for results and improvement of practice, but to support a wide range of other demands. They include: (a) supporting the development of a learning and evaluation culture; (b) monitoring and reporting on evaluation policy implementation and compliance; and (c) responding to the reform agendas for coherence, mutual accountability for results, national evaluation capacity development, gender equality and human rights, and diversity and inclusion. Those demands are reflected in several General Assembly resolutions on

---

[7] Regulations and Rules Governing Programme Planning, the Programme Aspects of the Budget, the Monitoring of Implementation and the Methods of Evaluation (ST/SGB/2000/8), annex, p.15.

[8] BusinessDictionary.com. Peter Senge, The Fifth Discipline, MIT, 1990.

evaluation[9] dating back to the 1980s and in a number of instruments concerning aid effectiveness, mutual accountability and results-based management.[10]

11. There are also the more recent changes in the global setting, including the imperatives of the post-2015 development agenda, which require transformations, restructuring and continuous evidence for fast action and innovation. They require cross-sectoral and integrative approaches and analyses of joint and system-wide work and global engagement. That has implications at three levels.

12. First, with the need for reforms and the importance of strategic direction setting, a significant number of central evaluation units are increasingly being asked to direct attention to evaluating the strategic repositioning and comparative added value of United Nations organizations, thus playing a greater role as agents of change. That has implications for changes in the types of evaluation being conducted, which are now of a more complex nature. Within a context of limited resources, some central functions are now focusing less on performance assessment or specific project evaluations, which are being shifted to the decentralized evaluation function embedded in management units. At issue is the capacity of the decentralized function to take on board the production of high-quality evaluations. Chapter III addresses the structural relationship between central and decentralized evaluation systems and the role being played by the central function to enhance the capacity and quality of decentralized evaluations and/or to control the quality of evaluation reports.

13. Second, the fast changes on national and global platforms highlight emergent and complex development processes in difficult environments. They require the use of new evaluation models that respond to: (a) the need for evaluation to see development not as a define and fix hypothesis, but as an endogenous process of transformation — development takes place in naturalistic settings and thus is not always controllable; (b) the fact that the best practice models that dominated the transfer of technology of development interventions are generally not viable in development contexts and that context-bound interventions are messy and difficult to evaluate; and (c) the need for just-in-time or real-time information on what works, why, and how, in order to support timely, relevant and fast-paced decision-making processes and mid-course correction all important for maximizing the relevance, effectiveness and value of interventions.

14. Third is the demand arising from the post-2015 development agenda for interdependence and interconnected ways of working together in the United Nations system. That would require going beyond coordination and engendering greater collaboration in cross-organizational programming and in the conduct of evaluations including examining the sum of the value and the impact of United Nations system response to cross-sectoral global, regional and country level needs and priorities. That is reflected in the more recent demand for the evaluation

---

[9] For example, General Assembly resolutions 59/250, 62/208 and 67/226.

[10] The Paris Declaration on Aid Effectiveness; the Accra Agenda for Action; the Busan Partnership for Effective Development Cooperation; and the Monterey Consensus of the International Conference on Financing for Development.

functions of organizations to participate with JIU in system-wide evaluation, such as via the policy for independent system-wide evaluation of operational activities for development of the United Nations system contained in a note by the Secretary-General (A/68/658-E/2014/7). It is also reflected in the Scaling Up Nutrition initiative which requires the United Nations to work in partnership with multiple actors in both programme development and in evaluation.

15. These demands have major implications for the evaluation function, requiring it to rationalize, restructure and address complementarities and alignments. That should be done among United Nations system internal evaluation functions. In order to enhance partnerships, alignments should also be done with the wide range of national and development partners, as well as private sector and other emerging global actors involved in evaluation.

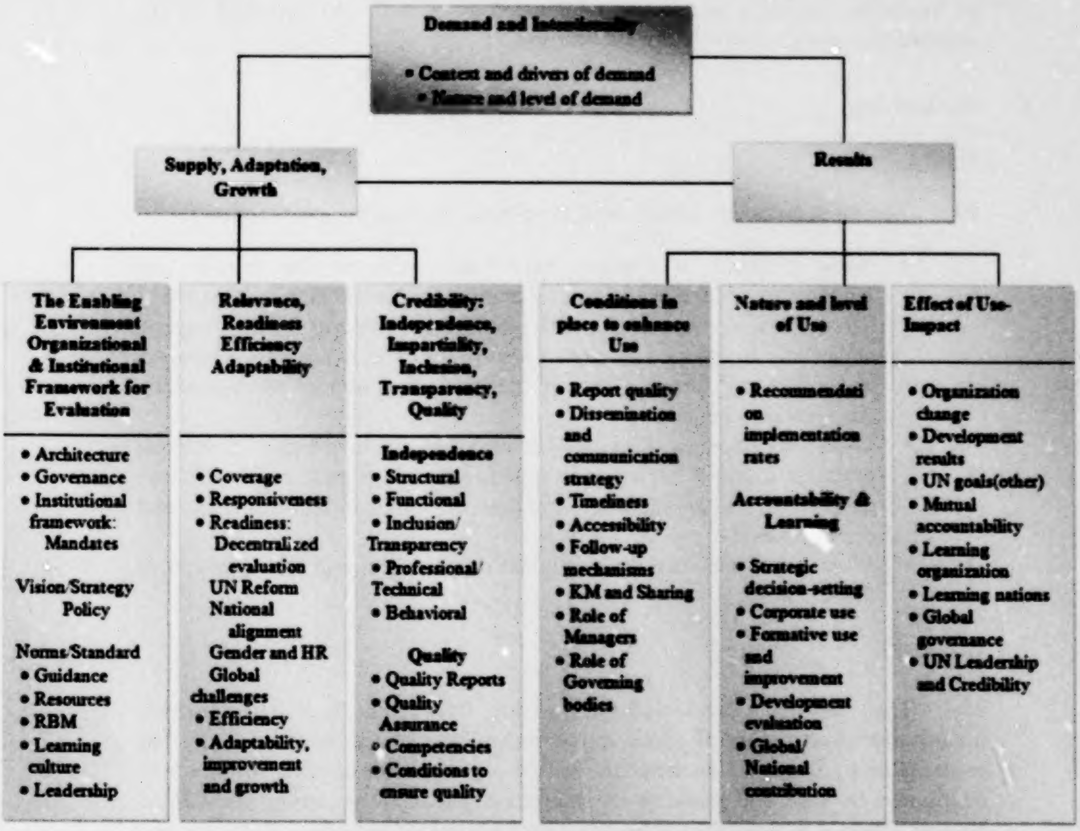
United Nations system response to the growing demand for and development of a high-quality evaluation function and support systems

16. In response to growing demands, a number of organizations in the United Nations system have sought over the years to develop a high-quality, effective and responsive evaluation function that supports accountability, learning and reforms. They have put in place business models, structures, policies, systems and mechanisms for the development of the core components of the evaluation function, of which the following are addressed in this study:

- Component 1. An adequate enabling environment and institutional framework that support the evaluation function
- Component 2. Independent and/or impartial evaluation to foster credibility
- Component 3. High-quality evaluation processes and products to engender credibility
- Component 4. The utility of the function, including use and impact of evaluation
- Component 5. The relevance of the system to respond to organizational demand for programmatic coverage and adaptability in responding to demand; and the related dimension of readiness to address demands associated with reforms and emerging global changes and challenges.

17. Figure I below sets out the conceptual framework of the evaluation function, highlighting the linkage between demand and key components in response to demand.

Figure I. Components of the evaluation function – JIU model


Demand and Intentionality
• Content and drivers of demand
• Nature and level of demand
Supply, Adaptation, Growth
Results
The Enabling Environment Organizational & Institutional Framework for Evaluation
• Architecture
• Governance
• Institutional framework: Mandates
Vision/Strategy Policy
Norms/Standard
• Guidance
• Resources
• RBM
• Learning culture
• Leadership
Relevance, Readiness Efficiency Adaptability
• Coverage
• Responsiveness
• Readiness: Decentralized evaluation
UN Reform
National alignment
Gender and HR
Global challenges
• Efficiency
• Adaptability, improvement and growth
Credibility: Independence, Impartiality, Inclusion, Transparency, Quality
Independence
• Structural
• Functional
• Inclusion/ Transparency
• Professional/ Technical
• Behavioral
Quality
• Quality Reports
• Quality Assurance
◦ Competencies
• Conditions to ensure quality
Conditions in place to enhance Use
• Report quality
• Dissemination and communication strategy
• Timeliness
• Accessibility
• Follow-up mechanisms
• KM and Sharing
• Role of Managers
• Role of Governing bodies
Nature and level of Use
• Recommendation implementation rates
Accountability & Learning
• Strategic decision-setting
• Corporate use
• Formative use and improvement
• Development evaluation
• Global/ National contribution
Effect of Use-Impact
• Organization change
• Development results
• UN goals(other)
• Mutual accountability
• Learning organization
• Learning nations
• Global governance
• UN Leadership and Credibility

18. When the results for the various components outlined above in paragraph 16 are attained, the expectation is that the evaluation function would: (a) enhance the accountability and oversight of the United Nations system on the basis of credible evidence; (b) enhance the quality of plans, strategies, policies, programmes and projects, as well as institutional reforms; and (c) enhance the leadership role and contribution of the United Nations system in sharing lessons learned from evaluation with national partners and on the regional and global landscape. Annex I provides an illustration of the logic model or results framework of the evaluation function supporting those goals and the ultimate effect or impact on the lives of people.

19. It is the structures, systems and mechanisms addressing those components that are the focus in this study. Annex II provides, for each of the five components of the function, a set of indicators that define what are identified as significant systems, operations and outputs that are important to enhance the function. Sixty-

six (66) indicators[11] were identified as critical for the function. They are comparable to bilateral and multilateral indicators. Details of the indicators and performance standards and methodology for assessment are provided in the supporting materials found in the JIU website.

C. Methodology

Main questions

20. The main questions raised about the evaluation function are as follows:

- What progress or changes have been made in the growth and development of the evaluation function in the United Nations system?
- What is the current level of development or maturity of the function according to the components identified to enhance its relevance, credibility and utility? What is the level of readiness for supporting changes and challenges?
- What variations exist among the different organizations? What good practices exist? What are the major strengths, systemic constraints and challenges? What alternative models are optimal for efficiency and effectiveness?
- What would enhance the continued development and advancement of the function?

Scope

21. This study was conducted system-wide. It covered 28 JIU participating organizations, consisting of funds, programmes, specialized agencies and other organizations. Annex III provides the list of organizations covered in the study configured by type and mandate. It should be noted that regarding the United Nations Secretariat, only the Office of Internal Oversight Services (OIOS) and the Central Evaluation Division is included. It did not include the departments within the Secretariat that have embedded evaluation functions. Information on those departments can be found in a recent OIOS report.[12]

22. The study focused primarily on the central function and secondarily on the decentralized evaluation function of the United Nations system. Annex IV (a) provides an overview of the architecture of evaluation in the United Nations system and how those two functions support different levels of corporate decision-making in many organizations. An extension of that architecture can be found in annex IV (b), which provides information on the departments of the United Nations Secretariat that have embedded evaluation functions under OIOS.

---

[11] The JIU maturity matrix has 70 indicators, but only 66 were assessed in the study. Two indicators did not yield valid assessments. Two are not assessed in a comprehensive manner for all organizations. Thus, four were dropped in the study.

[12] Activities of the Office of Internal Oversight Services on peacekeeping operations for the period 1 January to 31 December 2012 (A/67/297 (Part II)).

23. Given that there is considerable variation in the terminology used in United Nations system organizations to describe the elements of the evaluation system, it is important to further clarify the distinction between the two levels of the evaluation architecture of the United Nations system that were the focus of the study — the central function and the decentralized function. Annex V provides details on the two levels. The central-level function in the United Nations system operates independently of the control of line management and, in some cases, independently of executive management or the head of organization in the planning, management, conduct and reporting of evaluation. It generally supports the broad policy, programmatic and strategic decision-making of the organization, although in many cases the central function conducts project evaluations of varying types.

24. Decentralized evaluations are planned, managed and conducted outside the central evaluation unit. They are embedded within programme and management units that are responsible for the planning and implementation of projects and programmes. That includes projects and programmes at the technical, regional and country office levels or departments. The decentralized evaluation function thus generally operates as an integral part of programme management, addressing evaluation, monitoring and other related analytical activities to support decision-making associated with the various phases of the implementation of projects and programmes.

25. Decentralized evaluations are thus not structurally independent, as they come under the control of line management in the planning, management, conduct, approval and reporting of the evaluation. That applies to both the evaluations managed and conducted by programme staff at their discretion (described as self-evaluation in some organizations) and evaluations that are commissioned or contracted to external consultants (for which the final approval of the report rests with the programme managers). That also includes evaluations of the United Nations Development Assistance Framework (UNDAF). The system for UNDAF evaluation was not addressed in the study, as it is potentially part of a proposed pilot study of the independent system-wide evaluation of operational activities for development.[13] Evaluations that are planned at the decentralized level under the control of other credible independent bodies, and for which the final approval and issuing of the report is independent, can be assessed to be independent. Their credibility, however, depends on the transparency of the involvement of all the key stakeholders. Details of the archetypes of the decentralized system are provided in annex V (b) and discussed in chapter III.

26. Annex IV (a) which shows the architecture of the Evaluation function includes a third level - which is the United Nations inter-agency and system-wide evaluation. The present study did not include the function addressing system-wide evaluation, since a more recent study focusing on that aspect has been

---

[13] Policy for independent system-wide evaluation of operational activities for development of the United Nations system (A/68/658-E/2014/7).

commissioned by the Secretary-General.[14] Of the 28 JIU participating organizations, four do not have a central function, but carry out ad hoc evaluations. The study provided brief descriptions of the evaluation function in those four organizations. The focus on the central function throughout the present study is thus on the remaining 24 JIU participating organizations.

27. This study did not assess the impact of the evaluation function. A description of this can be found in the top part of the results framework of the evaluation function in annex I. It did, however, assess the systems for impact assessment. It also assessed the attainment of a selected number of the outcomes or proxy outcome indicators of the function listed in annex I.

28. **The objectives of the study** were to:

Objective 1. Provide information on the evolution of the evaluation function dating back to 1977. This provided a context for highlighting the function's progress and its shifting roles in response to demand. It is of historical importance and useful for tracking the changes in the function.

Objective 2. Describe the status and analyse the degree to which the central evaluation function of organizations meets established professional standards in the components of the evaluation function, and highlight cross-cutting strengths and weaknesses in those components. This provided information on variations between the organizations and, in so doing, highlighted common and different organizational challenges, as well as exemplary practices.

Objective 3. Present an analysis of the level of readiness of the evaluation function to adapt to global changes and meet global challenges and the emerging imperatives of the post-2015 development agenda, drawing on attributes identified by JIU and the heads of evaluation offices as important for responsiveness to change and the demands on the United Nations system.

Objective 4, Describe the status and analyse the development of the decentralized evaluation function via a rapid review. This objective was added to the study so as to provide a more comprehensive perspective of the UN evaluation system and hence provide a basis for a balanced and strategic approach for supporting the central and decentralised functions.

Objective 5. Identify structural and systemic constraints that should be addressed in order to enhance the capacity to deliver credible evaluative evidence and to enhance its use across the United Nations system.

Objective 6. Based on the conclusions reached and exemplary practices identified, make recommendations on how to enhance the continued development and relevance of the evaluation function for the United Nations system for all key players involved in the function – legislative bodies, management, and evaluation units.

---

[14] Angela Bester and Charles Lusthaus, *Independent System-Wide Evaluation Mechanism: comprehensive review of the existing institutional framework for system-wide evaluation of operational activities for development of the United Nations system mandated in General Assembly resolution 64/289* (Final report, April 2012).

Methods

29. The study looked at the status of evaluation in 2013, while also looking to the future with respect to what is needed in order for evaluation to play a major role in supporting the United Nations system achieve its goals and objectives. Both quantitative and qualitative methods were rigorously applied to ensure the accuracy, validity and reliability of the evidence used to reach conclusions. The analysis of the performance at the central level of 24 organizations was based on secondary analyses of over 500 documents (with an average of 20 documents per organization), cross-validation of data, interviews and consultations with over 150 staff and managers from the United Nations evaluation and management units, and interviews with Member States and external partners.

30. There was validation of concepts and constructs; involvement of key stakeholders; cross-validation and triangulation of evidence; extensive use of statistical tests of significance to assess the confidence that could be placed in the findings presented;[15] and assessment of rival hypotheses using qualitative methods. In accordance with General Assembly mandates, it addressed both the hard systems and mechanism of the evaluation function and the culture of evaluation and learning. The evaluation function has multiple players with complementary responsibilities — legislative bodies, executive heads and staff, and evaluation managers and staff. The assessment of the level of development of the evaluation function presented in this study therefore reflects the combined effort of all these players.

31. While the central evaluation and decentralized evaluation functions share the same five components that define the evaluation function, they have different priorities in their goals and they have different configurations, levels of development and standards that warrant different assessment methods. Thus, for the 24 central functions, performance on the components was assessed against a maturity matrix with a set of 66 indicators that were benchmarked against established standards endorsed by UNEG, JIU and development partners. The assessment at the central level on those indicators was based on five levels of performance (low, below average, average, above average and high). Paragraph 53 below provides details on the rating scale used for those levels.

32. For the decentralized evaluation function, the evaluation was challenged by the absence of agreed standards on the policies and systems an organization needs to have in place to support an efficient and effective decentralized evaluation system. Thus, the study first carried consultations and then established a conceptual framework fit for decentralized evaluation. That framework was then used to conduct a rapid assessment of the status of decentralized evaluation across the United Nations system. A questionnaire was sent to organizations that had

---

[15] To enhance rigour, the study used multiple statistical analyses appropriate for parametric and non-parametric statistics, including analysis of variance (ANOVA), the Kruskal-Wallis one-way analysis of variance and several post hoc tests. It is acknowledged that factorial, multiple regressions and other advanced statistics would have enhanced the explanatory power of the study. A more elaborate scope for the study in terms of work and resources would have accommodated such methods.

indicated that they had a defined decentralized function and to officials who had been identified as having oversight responsibility over the decentralized function.[16] That was complemented by extensive consultations, interviews with staff members and stakeholders and document analysis.

33. For the four organizations without a central function, a concerns-based questionnaire was administered to assess their current status, issues, systemic constraints and intentions with regard to developing a corporate evaluation function. The questionnaire was completed by three organizations. Information on all questionnaires used in the study and the methodology for the maturity matrix can all be found in the JIU website.

Challenges in assessment and in system-wide analyses

34. The study met with challenges with regard to system-wide measurement and analysis this arising from differences in mandates and structures of the various organizations, the non-comparability of data, and the differences in meaning applied by the organizations to the concepts in evaluation and its function. In order to address those challenges, various methods were applied. A definition of concepts or constructs was established and used throughout the study. Standardized and criterion-referenced approaches were used in assessment. The variations between different types of organizations were highlighted and possible explanations provided. Weights were applied and the data were calibrated for missing information. The use of inferential statistics for cross-comparisons was dropped when there were huge data gaps or total non-comparability; in those cases, descriptive analyses were provided. The study paid particular attention to validity of the assessment for small organizations and the decision to use the maturity matrix was based on consultations and analysis of its validity and value for assessing the evaluation function of a small organization. Data calibration for small organizations did not make a significant difference to the findings.

Limitations

- For the central function, the development and use of the maturity matrix occurred concurrently, but in the end all data sets were corrected and updated for consistency for 2013.[17]
- Interviews were planned with a range of stakeholders, but not all the interviews could be conducted, owing to time and budget limitations, as well as difficulties in gaining the required access in a timely fashion.
- Benchmarks for which the evidence was judged to be insufficiently robust because it could not be drawn from more than one reliable source, and benchmarks that were not deemed applicable with regard to the different

---

[16] The official identified as having that oversight responsibility varied. In some organizations, oversight was provided by the central evaluation office. In others, it was provided by a central support unit or unit in charge of results-based management (RBM).

[17] The exception was the final assessment for the United Nations Population Fund (UNFPA), which was conducted in February 2014. The evaluation function was undergoing restructuring in 2013, hence the need to finalize the assessment in early 2014. In addition, the data for UN-Habitat were not validated via interviews and information update.

mandates of organizations, were not included in the assessment and analysis.

- For the rapid review of the decentralized evaluation, the analysis was based on data from a questionnaire completed by officials with oversight responsibility, and data derived from interviews and document analyses.[18] A more in-depth data collection exercise would have taken samples from over 940 units/offices that manage and conduct decentralized evaluation, as well as a large number of staff, such as project managers, who conduct what is described as *self-evaluations*. There was not enough time or resources to do this form of rigorous data collection. The information from the questionnaires with oversight officers, interviews, and document analyses however provided an appropriate basis for a rapid review. Potentially, the series of reviews of RBM in the United Nations system, planned by the JIU in 2015, could more fully examine the decentralized evaluation system.
- The analysis of the evolution of the function over the 36-year time span covered in the study synthesized different data sets from various reports. The information for this is therefore more illustrative than absolute. The analysis was complemented by consultations and validation by various stakeholders who have been involved in the evaluation function over the past 15 years or more.
- Given the limitations in capacity, the study did not use advanced statistical methods that would have enhanced its explanatory power. It did, however, apply a sufficient number of descriptive and inferential statistics for a selected number of analyses.[19]
- The study identifies top performers on selected dimensions of the function and quality criteria. Explaining what drives high performance is done to a limited degree. More extensive conclusions would have required more data and analyses than were available. The study nevertheless includes many analyses of hypothesized relationships in supporting conjectures made.

D. Acknowledgements

35. The JIU Inspectors would like to thank all those who provided guidance, substantive comments and useful input. The scope, approach, instruments and analysis benefited from consultations with members of UNEG, who also provided extensive information for assessments. The study also benefited in conceptualization, validation and advice from an international expert and a pro bono, informal advisory body with the following membership: the former Executive Secretary of JIU, in her capacity as an evaluation expert; managers and

---

[18] Those data were limited to funds, programmes, specialized agencies and other United Nations entities. They did not concern the embedded departments of the United Nations Secretariat.

[19] The analytical methods used included Spearman's rho and Pearson's r correlations, t-test, ANOVA and Kruskal-Wallis and many post hoc tests of significant differences. Cross-tabulations with descriptive data were found to provide concrete information.

staff of various United Nations organizations; an evaluation manager from the Department for International Development of the United Kingdom of Great Britain and Northern Ireland; a representative of the Office of Institutional Assessment and Evaluation at Howard University in the United States of America; and from a regional expert on development. Consultations in the early stages of the project with the OIOS Inspection and Evaluation Division, the World Bank, the Development and Assistance Committee and the Development Co-operation Directorate of the Organisation for Economic Co-operation and Development (OECD DAC) were all important for benchmarking purposes. Last but not least, the Inspectors would like to thank the JIU focal points for their extensive support in coordination.

II. THE CENTRAL EVALUATION FUNCTION: EVOLUTION, GROWTH AND LEVEL OF DEVELOPMENT

A. Changes in development: purpose, structural arrangements, coverage and management

Conclusion 1: The central evaluation function has evolved and matured substantially in terms of systems operation, and its role, structures and standards have developed since the JIU reports of 1977 and 1990. Its role has diversified: it has continued to support the decentralized evaluation function, while also supporting broad and strategic corporate/central-level decision-making. Its role has also expanded to respond to varying demands. Organizations have remained pragmatic and cost-conscious in developing their evaluation systems and this applies to both the structural arrangements made and the business models designed for the management and conduct of evaluation. It is, however, evident that the resourcing of the central function has not changed much in real terms. There is a need for the United Nations system to reflect on how seriously it wants to know the impact it is having and how it could enhance its relevance on the national, regional and global landscape.

Changing role

36. Evaluation in the United Nations system has undergone substantial growth since 1977[20] when the United Nations system was characterized by a *self-evaluation* system focused on the need for learning and improvement[21] and the central evaluation offices had the primary role of "central oversight and co-ordinating to ensure quality and performance of the predominant system of self-evaluation".[22] By the time of this study, the role of the central evaluation office among the 24 JIU participating organizations examined in this study had diversified to focus on support to strategic and corporate-level decision-making. Many organizations continued to carry out in parallel self-evaluation or decentralized evaluation and a range of other analytic and evaluation related tasks to respond to demands. Table 1 below provides an illustration of this progression since 1977.

Structural arrangement

37. The location of evaluation units in the structure of the organizations has become more stable. In the early years, they were attached *to... "top executives in the organizations; integrated with planning and programming activities; or combined with administrative financial service units*".[23] Central evaluation units were characterized as "*grafted on to existing structures in a rather uncertain and insecure fashion, and have not yet really become established as an integral part of the organization's management process*".[24] What was significant was the small number of staff, with 60 professionals in evaluation system work during the

---

[20] See JIU systematic and comprehensive studies on the evaluation function between 1977 and 1985.

[21] This is reflected in the definition of evaluation at that time as a learning and action-oriented management tool (A/34/286, 6 June 1979).

[22] "Second report on evaluation in the United Nations system", JIU/REP/81/6, p.4.

[23] Ibid.

[24] Ibid.

period, representing less than 1 per cent (0.7 per cent) of total professional staff (18,500 staff for 23 organizations), and with very limited resources committed which were not commensurate with the number and scope of organizations' project and programme results.[25]

38. At the time of the present study, 24 of 28 JIU participating organizations have defined policies, well-defined governance and structural arrangements, and clear systems and business models for enhancing efficiency and effectiveness of evaluations conducted. They also have an extensive array of mechanisms to enhance the independence, coverage, credibility and utility of the evaluation function. The level of development of these mechanisms is the focus in this study.

39. The Inspectors observe that the organizations have remained pragmatic and cost-conscious in developing evaluation system and this applies to both the structural arrangements made by the various organizations and business models designed to use with efficiency the comparative advantages of staff and consultants in carrying out evaluations. It is, however, not very clear to the Inspectors whether the level of commitment and resourcing of evaluation has changed to a significant degree. The paragraphs and tables below highlight these issues.

Table 1: Phases of the evaluation function and changing roles of the central evaluation function

Period	Phase I (1977-1980)	Phase II (1980s)	Phase III (2000-2005)	Phase IV (2006-2013)	Phase V (2013 onwards)
Objectives and focus	Leaning Evaluation designed for use in projects	Learning Use in project and onset of quality enhancement by central unit	Accountability Focus on independence and improvement of programs and projects in meeting objectives	Balancing: Accountability. Independence. Quality. Learning. Coherence. National evaluation. & Other demands	Emerging Trends Balancing various demands and emerging changes and challenges
Predominant character of the evaluation system	Widespread project evaluations built into project implementation (*Self-evaluation later termed decentralized evaluation*)	Widespread project evaluations built into project implementation	Corporate central evaluation being established: Focus on putting in place system for central and decentralised functions.	Developments of systems at central level for independence, credibility and utility (based on UNEG Norms and Standards)	Strategic, routinized, complementary roles of central and decentralised evaluation systems within a defined strategic framework to support organizational reform and transformation
		Coordination and quality enhancement of decentralised evaluations by central evaluation unit	Decentralised evaluation with two main models: (i) decentralized system with institutional framework; (ii) ad hoc system done at discretion of staff and managers	Decentralised evaluation with two main models: (i) decentralized system with institutional framework; (ii) ad hoc system done at discretion of staff and managers	Alignments and collaboration for evaluation of system-wide value Alignments for national capacity development and global partnerships

---

[25] Ibid, para. 18.

40. Thus, for the 24 organizations with a central evaluation function supporting central-level decision-making, the choice of organizational location of the central function has varied: 10 organizations have well defined stand-alone evaluation functions; 14 are co-located, of which 9 are co-located together with the broader oversight function under the head of the office of oversight (either as a separate unit with a chief or fully integrated with the other functions of oversight); and 5 are co-located with executive management (and integrated with policy, planning, management or research) (see Table 2 below).

41. The choice of location is reportedly influenced by many factors tied in varying forms to hypothesized gains for independence, efficiency, relevance and utility. In four of the seven larger organizations (defined in budgetary terms of annual organizational expenditure above US$ 1 billion), the choice has tended towards creating independent self-standing units, partly in response to the UNEG norms and standards and Executive Board demand. The choice of co-location with other functions is by contrast predominant in small organizations (annual organization expenditure below US$ 300 million per year) and reflects consideration of either cost and efficiency or the priority given to: (a) establishing independence (by co-locating the evaluation function with other independent oversight functions), or (b) the perceived relevance to and direct impact on the organizational culture given to the function by management (by co-locating it with executive management). The study presents findings on the effect of the different structural locations on the development of the evaluation function.

42. Of the 28 organizations, 4 are without a central evaluation function — the International Telecommunication Union (ITU), the World Tourism Organization (UNWTO), the Universal Postal Union (UPU) and United Nations Office for Project Services (UNOPS). They provide various reasons for this situation, among which the following stand out: (a) evaluation not being of high priority owing to a lack of demand for evaluation by governing bodies and thus limited funding; (b) an organizational culture more oriented towards delivering activities than considering the results achieved; and (c) the overall modest budget of the organization.[26] There are also reported concerns that, given the type of organization including its capacity and the cultural context, evaluation would not be feasible or add value.

43. In the view of the Inspectors, the explanations do not fully justify the absence of an evaluation function. There is a need for these organizations to revisit the role of an evaluation function in enhancing the effectiveness, efficiency and sustainability of their work. At the time of completion of this report, ITU and UNWTO informed they have now initiated plans for developing an evaluation function. **The Inspectors submit that, given the circumstances, current reflections on the development of the institutional framework and policies for the evaluation function in small organizations be preceded by a well thought-out vision and strategy, including its added value, priorities, synergies and**

---

[26] The aggregate annual organizational expenditure for three of the four organizations is about US$ 196 million per annum. The budget range for the 28 participating organizations is between US$ 27 million and US$ 5.3 billion per annum.

alignments with other oversight and knowledge-development systems existing in the organization, as well as alignments with other United Nations agencies and external partners. The case of UNWTO utilizing the services of the UNOPS ethics officer as opposed to setting up its own office, is an example of a practical service delivery with efficiency.

Table 2: Structure of the central evaluation function: size and structural arrangement for 24 JIU participating organizations

Structural arrangement		**Size of organization by annual budget (US$)**		
		Large US$ 1–5.5 billion	**Medium US$ 300 million–1 billion**	**Small < US$ 300 million**
A, Stand-alone **N = 11**	**A1: Function operates as a completely separate office of evaluation**	UNDP WFP FAO	ILO UNFPA	UNIDO
	A2: Function operates as a separate office but in the office of the executive head	UNICEF	UNEP	UN-Habitat UN-Women UNODC
B. Co-located **N= 13**	**B1: Function located within the independent internal oversight office**			
	Function operates in separate unit with a chief or director, operating in parallel with other oversight functions: ***3 organizations***	Secretariat of the United Nations/ OIOS	UNESCO UNRWA	
	Function is integrated with other oversight functions ***6 organizations***	WHO[27]	IAEA WIPO	ICAO IMO WMO
	B2: Co-located with Executive management (policy, planning, research)			
	Function is fully integrated with policy, planning, management or research: ***4 organizations***	UNHCR	UNAIDS	UNCTAD ITC

---

[27] The data source here for the analysis is from 2013. Some recent structural changes include changes in WHO, where in August 2014 the organization moved the evaluation function from the responsibility of the Office of Internal Oversight Services to stand alone as part of a new function "Evaluation and Organizational Learning".

Core function — production of reports to support decision-making

44. The role of the central evaluation function focuses on conducting evaluations and producing evaluation reports as its core function to support decision-making on policies, programmes, projects and operations, while also carrying out a diverse range of evaluation-related tasks to support the institutionalization of the function. During 2012–2013, the central evaluation offices of the 24 organizations produced 584 evaluation reports of various types, of which: 26 per cent were complex, addressing thematic, country, regional and global programmes; 4 per cent were joint evaluations; and 64 per cent were small and large-project evaluations. The latter project evaluations are done mainly by small organizations and those with a predominantly normative mandate. In general, for seven large organizations and those with a defined decentralized structural arrangement,[28] project evaluations are done principally at the decentralized evaluation-level and over 636 evaluations are produced on a yearly basis.[29] Given their number, the reports produced at central and decentralized levels represent a large knowledge pool for the United Nations system.[30]

45. This knowledge asset of the United Nations system is, however, dispersed throughout the offices and websites of United Nations system organizations. After years of debate, UNEG established in September 2014 a public repository for all evaluation reports of its members and the central offices. The repository contains more than 4,800 reports dating back to 1997, with contributions at this stage from 11 UNEG members.[31] This database can significantly enhance the sharing of knowledge and represents a major effort to facilitate cross-cutting analyses and the learning of lessons for system-wide use. A recent survey on the changes and challenges of the United Nations system's work in development recommends a single gateway to the United Nations knowledge assets and publications, given that "the UN's Internet presence is notoriously hard to navigate and poorly indexed for outsiders".[32] This UNEG effort represents one step towards responding to this demand.

---

[28] UNESCO, WFP, UN-Women, WHO, UNFPA, UNICEF and UNDP.

[29] This is evidence from the data provided on the decentralized evaluations of the seven organizations.

[30] If the reports coming from the embedded departments of the Secretariat of the United Nations were added, the number would increase tremendously.

[31] See www.uneval.org/evaluation/reports. Users can search for reports by agency, type of evaluation, country, region or keyword.

[32] Barbara Crossette, "Is a Better United Nations system possible? Global Experts Respond to an Independent Survey", 26 May 2014. Available from http://passblue.com/2014/05/26/is-a-better-un-system-possible-global-experts-respond-to-an-independent-survey/.

46. In most cases, the choice of what to evaluate is based on a combination of well-defined strategic criteria.[33] It is reported that coverage is not adequate to respond to the demands from within the organization. The main constraint on adequate coverage is noted to be inadequate resources. There is also the need at this time to prioritize use of existing resources to enhance the quality of evaluations, thereby limiting risks of poor credibility.

Business models

47. In carrying out these evaluations, the 24 central units have used a mix of different business models. While in the earlier years, the central function commissioned[34] a large number of evaluations to be conducted by external independent consultants, at the time of this review, central functions have full control and are totally accountable for all evaluation reports. There are a few exceptions. The main model used is one where staff takes a leadership role in managing[35] and in ensuring adherence to quality standards and the institutional accuracy and value of evaluation when working with expert consultants in conducting the evaluation. The business models used for managing the conduct of the evaluation in the various organizations differ because of the need to enhance the efficient use of resources and improve quality, while also ensuring impartiality (see annex VI).

Staffing and consultants

48. The 584 evaluations conducted by central functions for the two-year period of 2012–2013 were managed and conducted by 160 full-time-equivalent professional staff and managers in the central units, which represents about 0.18 per cent of staff in the organization.[36] The evaluations are conducted using the services of 989 expert consultants, who are estimated to provide the equivalent of 32 full-time positions. Thus, staff and consultants (making up approximately 192 full-time-equivalent positions[37]) all together make up less than 0.3 per cent of

---

[33] The organizations use a combination of criteria for selecting evaluations to be conducted, of which the predominant ones are: (a) linkage with organizational strategy (for 20 organizations); (b) internal demand by management (for 20 organizations); (c) external demand by donors, generally when there is a protocol agreement (17 organizations); (d) funding amount (for 14 organizations); and (e) evaluability of the activity (12 organizations). Other criteria applied included; risk assessment (10), consideration of emerging global trends (9), coherence with other United Nations system organization plans (8), and the potential for scaling up the activity based on evaluation (4).

[34] Commissioned evaluations are under the control of external bodies contracted to conduct and report on the evaluation. The prevalence of this form in the past is associated with the lack of independence of some of the evaluation functions.

[35] This includes developing terms of reference, contracting consultants, managing contracts, managing the conduct of the evaluation with due regard to organizational and institutional set-up, as well as standards and guidelines of the evaluation units and of UNEG, and quality assurance.

[36] Analysis is based on complete data available for 17 participating organizations. The total staffing cost is about US$ 534,711,127 for 2012–2013 for 160 staff members. Total expenditure for 989 expert consultants is US$ 21,856,019.

[37] The full-time equivalent staff and consultants makes up less than 0.3 per cent of organizational staff. There are no benchmarks against which to assess this ratio. It is, however, smaller than the 0.7 per cent ratio of 1981, as indicated in paragraph 36 above.

organizational staff,[38] a ratio well below the 0.7 per cent reported in the 1980s. On average across the United Nation system, each staff member (plus consultant) produces about three evaluation reports of different types every two years. These figures are estimates based on the available information, which is not always stable and require caution in their interpretation. The latter point of caution is important given the sporadic evidence which indicates that evaluation staff members actually work more than 8 hours daily.

49. Staff dedicate on average 24 per cent of their time to other evaluation-related or other oversight, research and strategic management activities and providing support to decentralized evaluations besides conducting evaluation. Annex VII provides information on the range of tasks taken on by staff in responding to demand. Organizations vary in the type and scope of work on non-evaluation report activities. The information in that annex is important in highlighting that an analysis of the impact and efficiency of the central evaluation function, as well as the financing of the function, would need to consider this range of activities that enhance the development of the evaluation culture in organizations in addition to producing evaluation reports.

50. While progress has been made in some areas, it is uneven in others and more needs to be done for the development and continued progression and advancement of the function, with due regard given to the role it could play in enhancing the effectiveness and sustainability of the United Nations system. The following sections highlight levels of development of the function and recommendations for improvement. Critical in this regard are more integrative approaches, operating from a well-thought-out strategic approach, balancing priorities and positioning of the evaluation function to enhance its added value and role in supporting the United Nations system currently "at the crossroads of development".[39]

B. Organizational level performance: patterns among and variations between organizations

Conclusion 2: Across all organizations, the central evaluation function in the United Nations system has made considerable strides in its overall level of development in meeting standards established for the core components of a high-impact evaluation function (i.e. *the institutional framework; relevance and readiness for emerging challenges; independence; quality; and utility*). Nevertheless, there are significant variations among the organizations in terms of the level and nature of development of these components. Based on the level and pace of development across the different components, organizations can be categorized into four clusters of maturity or development level.

---

[38] This is a gross estimate based on available numbers. Calculation is based on the 2012 data for 52,915 staff members (not including the departments of the Secretariat since they were not included in the study).

[39] Bruce Jenks and Bruce Jones, "United Nations Development at a Crossroads" (2013, Center on International Cooperation, New York University).

51. A historical analysis of the evolution of the central function in the United Nations system dating back to 1977 shows progression in level of maturity of the development of systems operation: (a) whether the systems for the function were being developed and were still rudimentary; (b) were developed but ad hoc or were being integrated; or (c) whether they were well-integrated and advancing to higher levels of systems operation, as well as having a broader strategic and outward perspective in enhancing the quality, relevance, and impact of the evaluation function.

52. The trend of the data, as illustrated in annex VIII and by anecdotal evidence from various United Nations system officials who have been involved in the development of the central function, suggests that the speed and magnitude of development have had ups and downs, and that there were times when the function was stagnant. The development of the central function appears to have picked up in the mid-2000s, which was linked to: (a) a major push for accountability regarding results and the value or difference made by the United Nations system; (b) JIU reviews of governance and oversight and recommendations for the establishment of evaluation functions, no matter how small the organization; and (c) the development of an evaluation profession in the United Nations system spearheaded by UNEG and the promulgation of its norms and standards in 2005.[40] These norms and standards provide a systematic institutional framework for the development of and oversight over the evaluation function. They also provided input to the standards used to guide the comprehensive analysis of the evaluation function carried out in the present study.[41]

53. For the current level of development, the assessment of the components was based on the ratings scale provided in the table 3 below with progression assessed on a continuum of five possible levels of maturity in carrying out the function in the five components. Details on the assessment method can be found in the JIU maturity matrix in the supporting materials in the JIU website.

---

[40] See www.uneval.org.

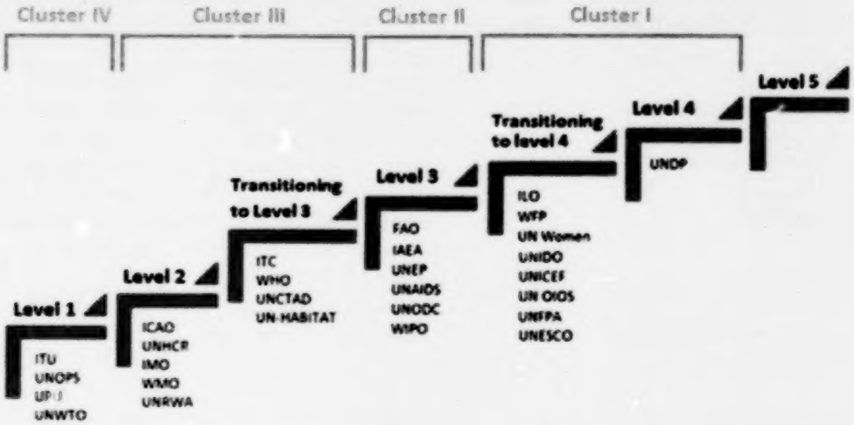
[41] The other standards used are those also used in benchmarking frameworks of multilateral organizations when assessing organizations of the United Nations system.

Table 3: Levels of development of the central evaluation function (maturity levels)

Maturity level or level of development	Score Range	Description of patterns in achieving established benchmarks*
Level 1 **Rudimentary** **VERY LOW**	1-2	Reflects non-existence or rudimentary levels on the components of the system: at best the organization is finding out how to establish a function and the expected results, systems, mechanisms and management structure needed.
Level 2 **Ad hoc** ***LOW***	3-4	The system for evaluation is ad hoc: there exist some of the elements of the various components of the function but these are not fully coherent or supported by a well-defined institutional framework.
Level 3 **Systematized and Routinized** ***AVERAGE****	5-6	The function is quite well defined: key measures and mechanisms of the various components are in place and operation is no longer ad hoc but has become routinized with some level of stability. The focus is on enhancing the integration, quality and institutionalization of the elements. As such, the orientation of the function is predominantly internally focused.
Level 4 **Refinement and enhancing use** ***HIGH***	7-8	Integration and institutionalization of the elements of the components of the function has mostly been achieved and the focus for improvement has shifted to enhancing the value added effect for the organization, as well as for the United Nations system as whole. Increased attention is given to partnerships and alternative ways of enhancing effectiveness, efficiency, sustainability, and impact.
Level 5 **Renewal** ***VERY HIGH***	9-10	At this stage, the focus is on the function as a major agent of change at organizational, United Nations system, national, regional or global level. Focus is also on making a contribution to cutting edge methods for evaluation and further enhancing the value of evaluation at all levels in an integrated fashion (organization, United Nations System, globally and nationally).

54. Figure II below shows how the 28 JIU participating organizations fall into four different clusters when judged in terms of the "overall development" of their individual evaluation functions across all components of the function.

Figure II. Overall level of maturity of the central evaluation function for 28 JIU participating organizations


Cluster IV
Cluster III
Cluster II
Cluster I
Level 5
Level 4
UNDP
Transitioning to level 4
ILO
WFP
UN Women
UNIDO
UNICEF
UN OIOS
UNFPA
UNESCO
Level 3
FAO
IAEA
UNEP
UNAIDS
UNODC
WIPO
Transitioning to Level 3
ITC
WHO
UNCTAD
UN-HABITAT
Level 2
ICAO
UNHCR
IMO
WMO
UNRWA
Level 1
ITU
UNOPS
UPU
UNWTO

55. In summarizing the status of organizations within each cluster in 2013, major highlights and distinguishing features include the following:

- *Cluster I* contains nine organizations where most of the expected systems and policies supporting the central evaluation function are in place and well-institutionalized. They are assessed as having an average of 65 per cent success level in meeting the established benchmark for the components of the function. Their challenge is primarily how to enhance the use and impact of the evaluation products and to balance increasing demands, including by playing a key role in addressing emerging organizational and global challenges. Organizations in the cluster are primarily funds and programmes (five), have stand-alone central functions (seven), and are large or medium-sized organizations (six). It is noted that the level of development of the funds and programmes meets expectations, given that they have had a long history in development work with donors requesting evaluations of projects that are funded. It is, however, important to note that statistical analysis shows no significant difference in the level of development between funds and programmes and specialized agencies. Organizations in this cluster dedicate substantial resources to evaluation and have many professional staff at higher grade levels

- *Cluster II* contains six organizations and reflects an average level of development of the evaluation function. A large number of the expected systems and policies supporting the central evaluation function are in place, but not all are fully operationalized. As such, the main focus of the central function is still on getting them operational to a stable level of development rather than on their use or impact, or on ensuring an outward-looking focus to enhance the added value of evaluation. Altogether, the organizations in this cluster have an average level of performance of meeting 56 per cent of the expected benchmarks for the study for all its five components. This group does also dedicate substantial resources to evaluation. It should be noted that compared with Cluster 1, it uses a large number of consultants to carry out the work.

- *Cluster III* has nine organizations in which the full complement of systems and policies expected are not yet in place and those which are in place do not operate to their full capacity or in an integrated manner. Organizations in this cluster perform below the average level of maturity. This cluster has a predominance of: (a) organizations that are small (with organizational budgets below US$ 300 million) and are co-located with audit units, with oversight units or with management units addressing policy and planning; and (b) large or medium-sized organizations that are also co-located and have the same challenges as the small organizations. This cluster has an average level of performance of 40 per cent, or below the average range, and has a long way to go in developing the function. Organizations in this cluster do not have, on a comparative basis, a substantial number of staff in the high professional grades (P4-D2) to carry out evaluations.

- *Cluster IV* includes UPU, UNWTO, ITU and UNOPS. A corporate central evaluation function *has not yet been established* although some rudimentary forms of evaluation exist. As noted in paragraph 43 above, ITU and UNWTO are initiating processes for developing a defined function.

For additional information on variations among the organizations, please see the supporting materials in the JIU website.

Conclusion 3: The evidence shows that evaluation functions that are co-located with oversight or are integrated within executive management units are in general lagging behind in their level of development. This is particularly prevalent among small organizations, where resource constraints in the organization are a determining factor in the co-location of the function. There are exceptions, however, to both co-location and size constraints. Success criteria include a governance and management structure that understands the strategic and added value of evaluation relative to other oversight functions, the provision of adequate resources and the quality of evaluations being ensured.

56. The statistical analysis of the 24 organizations with a central evaluation function shows that stand-alone functions are more developed than co-located functions; large and medium-sized organizations are more developed than small organizations. Thus, when looking at the *nine organizations in cluster III with average or low performance;* all nine have co-located central evaluation functions: five with management and four with other oversight. Six are small (annual budget

below US$ 300 million per year). The effect of this interaction is illustrated in table 4 below.

Table 4: Level of development of the evaluation function by size of the organizations' overall annual budgets and the location of the central evaluation function

	Small		Medium		Large	
	Organization	*Grade*	*Organization*	*Grade*	*Organization*	*Grade*
Stand alone	UNODC	5.2	ILO	6.7	WFP	6.7
	UNIDO	6.4	UNEP	5.8	UNDP	7.1
	UN Women	6.4	UNFPA	6	UNICEF	6.3
	UN-Habitat	4.2			FAO	5.9
Co-located/with management	ITC	4.9	UNAIDS	5.3	UNHCR	3.8
	UNCTAD	4.4				
Co-located with oversight/audit	IMO	3.7	UNESCO	6.1	UN-OIOS	6.2
	WMO	3.5	UNRWA	3.4	WHO	4.7
	ICAO	3.9	WIPO	5.2		
			IAEA	5.9		

3.0-4.9 = Below average and low

5.0-6.0 = Average

6.1-8.0 = Above average and high

57. It is, however, important to note that there are *pockets of excellence* suggesting that neither factor — size or co-location — *precludes* the development of a high-level function, as there are exceptions in both cases. When small organizations invest in the function (with good budgetary support, a good number of evaluation professional-level staff and of a higher grade, and a clear vision of its role and its added value and support of the organization), as in the case of the United Nations Entity for Gender Equality and the Empowerment of Women (UN-Women), the United Nations Industrial Development Organization (UNIDO) and the United Nations Office on Drugs and Crime (UNODC) found in clusters I and II, the level of development is above average. Likewise, large and medium-sized organizations or entities such as OIOS and the United Nations Educational, Scientific and Cultural Organization (UNESCO), where the function is co-located but functioning separately from the co-located oversight functions, are reasonably well-staffed or resourced and perform at an above-average level, further emphasizes the point that being merely co-located is not the determinant of low performance.

58. The analysis of explanatory factors indicates that at issue with functions that are co-located within oversight offices in some organizations, is the perceived loss of identify of the evaluation function. The analysis indicates that this has the effect of limiting the added value of the evaluation function including its resilience and flexibility in addressing change and enhancing the culture of evaluation and the learning organization. Also, in a number of cases of small and/or co-located functions, the mechanisms for quality assurance and quality enhancement, such as peer reviews and the assessment of the quality of reports, is weak, possibly reflecting low investment in the function. Likewise the visibility of the function and its added value is limited or may not be fully ensured, since these organizations do not typically present and discuss a full annual report on evaluation to their legislative bodies. Statistical analysis indicated that this discussion with legislative bodies is an activity found to be most significant in ensuring the visibility and development of the function. In most cases, information on evaluation is embedded in the annual report with information on other oversight activities. These factors should be taken into consideration to enhance the capacity and visibility of the evaluation function in all organizations in which the evaluation function is co-located.

59. Contrary to expectation, analysis shows that the evaluation functions integrated with other management functions such as strategic planning, research and general management do not have a higher level of use of evaluation or have a better-established learning organization. These two features are often hypothesized to bring added value to such co-located functions. However, such co-located organizations do not necessarily score at a significantly higher level on having traction or significant gain within the organization in developing the evaluation culture or learning organization. Analysis of the pattern of strengths and weaknesses of such functions do not reveal any discernible pattern of coherence or prioritization, suggesting the existence of an ad hoc approach to the evaluation function.

60. The implementation of the recommendation below will enhance the effectiveness of the function.

Recommendation 1

The executive heads of United Nations system organizations in which the central evaluation function is co-located and integrated with other oversight functions or integrated with the executive management functions should enhance the function and ensure its quality, integrity, visibility and added value.

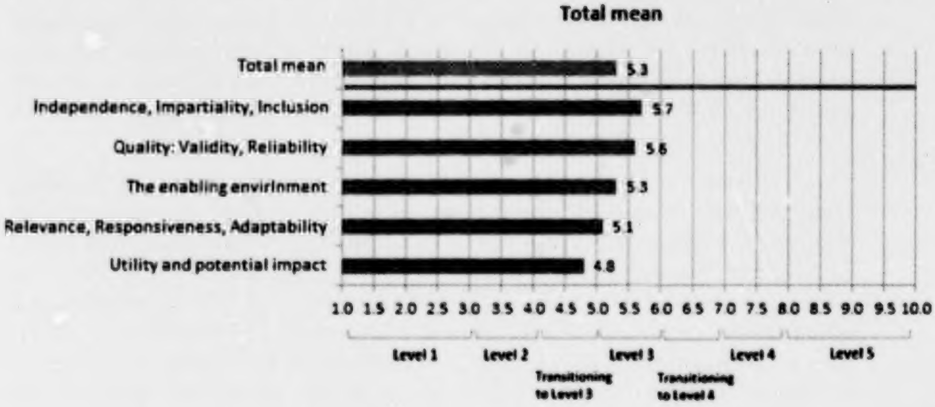
C. Components of the function: strengths, challenges and systemic constraints

Conclusion 4: The level of development among four components of the function (*enabling environment, relevance, independence, quality*) across 24 organizations is relatively at the same level and they all have been ranked "average" in the maturity assessment scale. However, this is not

the case for the fifth component: the utility of the function with a focus on the use of evaluation. For this component, the level of development is significantly less advanced and below average on the assessment scale. In general, within each component, both strengths and weaknesses are observed. The pattern of strengths and weaknesses oftentimes reflect imbalances within the system, systemic constraints, and variation in value systems or risk factors that may be common across the organizations.

Figure III below reflects the above conclusion and shows the level of development of the five components of the function.

Figure III. Level of development and mean score of each of five evaluation components



61. Based on the analysis of findings on the strengths and weaknesses of each of the five components, the following observations and conclusions are provided for each component of the function.

Component 1: The enabling environment, institutional framework and support systems

62. An earlier JIU report already stated that "*...even the well intentioned and carefully designed evaluation systems cannot succeed without the demand and the firm support from the governing and legislative body and from management and staff throughout the organization. Nor can it survive without a defined institutional framework of policies, norms, rules of the game, and standards*".[42]

63. Table 5 shows strengths and weaknesses according to the indicators assessed for this component. Analysis of some of the major patterns from the results for this component and the conclusions made are summarized below.

---

[42] Status of internal evaluation in organizations of the United Nations system, JIU/REP/85/10.

Table 5: The enabling environment — strengths and challenges

Areas where progress has been made	Areas in which progress is lagging
▪ Mandates/resolutions for evaluation from governing and legislative bodies ▪ Governance structure with accountabilities and responsibilities defined ▪ Role of member States/legislative body (accountability driven) ▪ Evaluation policy covering UNEG norms and standards ▪ Periodic review of policy implementation and revision in policy ▪ Use of non-core resources to support programme delivery ▪ UNEG norms and standards and their adoption and adaptation	▪ Adequacy of resources ▪ Defined normative framework to guide resource allocation ▪ Senior management leadership role and value of evaluation and promotion in organization ▪ Organizational culture for results, accountability, evaluation and learning fully rooted in organization ▪ Vision and strategy grounded on organization ▪ Alignment with other knowledge, oversight/audit and inspection, and decision-support systems — strategic alignments are critical.

Role of member States, mandates and polices

Conclusion 5: Continued strong demand appears to be a key reason for the increased number of organizations that have institutionalized evaluation systems and policies. This sustained demand has mostly come from the governing bodies. It has, however, led to evaluation functions that are primarily focused on meeting accountability demands. The establishment of evaluation polices based on UNEG norms and standards have also enhanced development. Such norms and standards also have a predominant accountability-driven structure.

64. General Assembly resolutions and organizational mandates and policies have played a key role in developing the institutional framework for evaluation. A large number of resolutions have been passed, dating back to 1980s. Between 2000 and 2013, three resolutions have been passed on the evaluation function.[43] Member States have played an important role both in guiding evaluation policy formulation and policy implementation. For organizations where member States involvement is less than robust, analysis indicates a lower level of development of the evaluation function. Member States have, however, focused more on accountability for results and the use of resources as opposed to other goals of the evaluation function. This focus on accountability is most important for meeting the demands for aid effectiveness and for maximizing the impact of aid but it needs to be better balanced with greater understanding of key features of the evaluation function that enhance development effectiveness (see also next section – Conclusion 6).

65. The institutionalization of the function has been significantly influenced by the introduction of the UNEG norms and standards in 2005, which resulted in most organizations aligning their policies accordingly. Two thirds of the

---

[43] A/58/587, A/RES/59/250, A/59/488Add.1, A/62/424/Add.2, A/RES/62/208/19, A/RES/67/226, A/RES/67/226 (27 of 30), A/RES/67/226 (28 of 30).

organizations of the United Nations system have an evaluation policy that seeks to be aligned with UNEG norms and standards, while the rest operate under a generic organizational accountability framework. The evaluation policies have played a significant role in defining the framework for the function: what it does, roles, responsibilities and accountabilities, types of evaluations and structures for evaluation. It should be noted that UNEG norms and standards also have a predominant accountability-driven structure.

66. The establishment of evaluation policies have also play a key role in the development of the function. The analysis of the content of existing evaluation policies shows that organizations have included to an extensive degree in such policies the UNEG norms and standards. This is not surprising given that all 24 organizations with central units are members of UNEG. Notwithstanding the above, only 46 per cent of these organizations were found to have a policy that describes in great detail the adaptation of UNEG norms and standards and inclusion of other norms to fit the context of the organization. However, policies that are not fully anchored in organizational priorities and the types of challenges to be addressed by the organizations' goals are likely to be less relevant. Missing from the majority of organizations, except three (the International Labour Organization (ILO), UNESCO, UN-Women), is an organizational strategy for the evaluation function and the use of that strategy to guide policy formulation. **It is recommended that the executive heads of the organizations of the United Nations system ensure the development of a well-defined, coherent and comprehensive strategy that is anchored in the organization and its plans for reform in the current context. This includes small organizations for whom alignments within and across organizations are of great importance to enhance efficiency while also retaining the professional integrity of the evaluation function and its added value for the organization.**

Learning culture and leadership

Conclusion 6: In responding to governing bodies, most organizations of the United Nations system have focused on evaluation mainly as an effective mechanism for supporting organizational accountability. In contrast, there has been no equivalent focus on ensuring that the evaluation function is integrated into the fabric of the organization and that it plays a role in its development, and that evaluation is used to strategically position the organization to respond to a constantly changing environment. Addressing this aspect of evaluation is critical to the sustainability of the function, particularly given the fast pace of change, emergent demands and complex challenges. The development of an evaluation function that is skewed towards accountability and away from the development of the learning organization limits the value and sustainability of the function.

67. While an accountability-driven culture continues to be most important, the changing global dynamics and emergent requirements of the global landscape, including the emerging demands coming from the on-going debates on the post-

2015 development agenda, call for an increased focus on learning and development of the learning and evaluation culture (i.e. the learning organization).

68. The definition of the learning organization in paragraph 9 above, as well as views given during consultations, guided the development of the rubrics (i.e. criteria, standards and levels of development) in the maturity matrix. Annex IX lists the indicators of the learning organization and evaluation culture for which information was compiled. The rubrics or performance standard used in the assessment can be found in JIU maturity matrix of the central function in the supporting materials on the JIU website.

69. The assessment of the learning organization examined: (a) how well the results, learning and accountability culture is rooted in the organization itself; (b) use of evaluation by those who could benefit most from the evaluation (double-loop learning); (c) sharing of information and development of communities of practice around evidence-based information; (d) adaptability and continuous transformation and self-renewal of the function and being creative and generative in producing knowledge and making a contribution to advancing development evaluation; and (e) making strategic linkages and alignments with other oversight and knowledge-generation functions.

70. The analysis, presented in annex IX, shows that the United Nations system performs "below average" level on rubrics that define a learning organization, including where the results and evaluation culture were embedded in the fabric of the organization. The analysis does show that the greatest strengths in the United Nations system development of a learning organization are: (a) the sharing of knowledge and development of communities of practice; (b) the continuous assessment and adaptation of the function; and (c) efforts directed at making a contribution to advancing development evaluation. The weakest areas of performance are in: (a) making strategic linkages with related functions (other oversight and knowledge management systems); (b) outreach and support to the decentralized evaluation function or national or other partner system; and (c) the use of evaluation for decisions and action. These areas are discussed in sections below addressing the conditions important for the impact of the function for change and challenges, and in the chapter on the decentralized function.

71. It is important to note that a number of organizations have in recent years focused on the development of the learning organization; these include the United Nations Environment Programme (UNEP), the World Health Organization (WHO) and ILO. WHO has carried out major restructuring of its evaluation function and, in August of 2014, it established a stand-alone evaluation unit as part of a new function for "evaluation and organizational learning".

72. Critical for the development of the learning organization are the leadership of the organization and the development of an incentive system that values innovation and risk management to prevent the practice of using timid and fail-safe approaches that do not sufficiently address challenges. The success of the evaluation function requires strong leadership important for direction-setting and for managing risks from innovations. The evidence shows that the performance of

the organizations on key leadership attributes and practices is not very strong. On average, only four organizations manifest high levels of performance across all the criteria used to assess leadership. The leadership of organizations was assessed to include the five JIU attributes of leadership[44] (see annex X). The analysis of strengths and weaknesses reveals that leadership role is above average in addressing systems for accountability and independence. Some 11 organizations manifest high levels of performance on this criterion. The institutionalization of evaluation and the development of the vision and strategy for the function is less strong. The implementation of the recommendation below will enhance the development of a strategic approach and the effectiveness of the function.

Recommendation 2

The executive heads of United Nations system organizations should adopt a balanced approach in addressing the purpose of evaluation for accountability, and for the development of a learning organization that has the appropriate incentive systems for innovation, risk-taking and the use of multidisciplinary perspectives.

Resources, financing and appropriate norms for budget allocation

Conclusion 7: Sustaining the central evaluation function's level of maturity and ensuring adequate coverage of issues to be evaluated balanced with adequate evaluation quality and impact is dependent upon the provision of resources that are adequate, predictable and well-managed. Evaluation functions in many organizations are under-resourced, especially in small organizations; and financing is unstable and unpredictable. There is generally an absence of a normative framework to guide and safeguard budget allocation. Only three organizations have established budget norms laid out in their evaluation policies, but the basis for the norms appears to be ad hoc and not based on a solid cost and a sound financial analytical foundation. Defining what this analytic framework should be has not been a priority, partly owing to a number of difficulties, including the fact that the budget structures of the evaluation function do not address the full scope of the function nor do they allow comparison or benchmarking across organizations.

Inadequate and unpredictable resources

73. Based on reported information from heads of evaluation offices and a preliminary analysis of what it would take to respond to demands, it can be concluded that the central evaluation function of most United Nations system organizations is largely under-resourced. It operates on average with 0.3 per cent of organizational expenditure. The analysis indicates that a 0.3 per cent investment

---

[44] JIU leadership attributes leadership by values and example, by information and communication, by guidance and discipline, by participation and by engagement in the use of evaluation.

is not adequate for organizations to operate a high-quality function to enable the United Nations system to understand the difference it makes in the world and in the lives of the people whom it is entrusted to help.

74. With the exception of two small organizations (UN-Women with 1.3 per cent in 2013 for both central and decentralized functions, and UNIDO with 0.5 per cent allocation for the central level), most other organizations have funding levels ranging between 0.05 per cent and 0.39 per cent. The statistical analyses in this study argue for the importance of adequate financing, as they show a high correlation between assessed levels of adequacy of funding and the assessed level of maturity of the evaluation function. Also, high-performing organizations in cluster I and stand-alone organizations dedicate in relative terms more resources to evaluation.

75. For 13 organizations with adequate data sets for analysis, non-core resources account for 24 per cent of resources for this group, albeit with major variations among the organizations. It is reported that non-core resources have greatly helped the delivery of evaluation programmes. Yet, such resources are also unpredictable, incur high transaction costs and affect the stability, relevance of coverage, efficiency and independence of the function. The analysis, although done with insufficient data to draw a firm conclusion, indicates that organizations with a high level of -core funding tend to have a higher level of development, although this is not statistically significant. They also tend to have strong measures in place to safeguard independence, limit transaction cost and effectively use the non-core resources to support delivery. Also, they manifest a higher level of independence compared to organizations with higher levels of non-core resources (see annex XI).

Budget norms and the basis for establishing norms

76. Protecting budget allocation for evaluation in general in the current context of budget cuts and organizational reforms is a concern among key stakeholders. There is generally an absence of norms and a normative or value-based framework to guide budget allocation. It is generally the view that providing budget norms helps to highlight the value of the function and to safeguard its viability. Four organizations have defined norms for budget allocation (UN-Women, the United Nations Population Fund (UNFPA), UNICEF, and WHO). They use different parameters in formulating their norms; however, the actual financial basis for defining the norms is not apparent.[45]

---

[45] For UN-Women, the target is set at 3 per cent of programme budget for both central and decentralized function; for UNFPA it is set up to 3 per cent of the total programme budget for the evaluation function both central and decentralized. For WHO, it is set between 3–5 per cent of the programme budget for both central and decentralized functions. For UNICEF, the policy calls for 1% of programme funds to be allocated to evaluation. The fact that the three organizations WHO, UNFPA, UN-Women that are organizations of different size all have an identical cost norm raises questions about the bases used for defining the norms. They nevertheless indicate an effort to define the value of the evaluation function in the organization and to safeguard its viability.

77. The different ways of developing budget norms is highlighted by the heads of evaluation units as a problem in benchmarking and standardization across the United Nations system. First, developing norms against programme budget assumes that the function only produces evaluations of programmes when, in addition to that, the evaluation function in the UN system carries out a wide range of other evaluation-related activities important to the organization. Second, evaluations conducted today are more complex than the typical performance assessments of the past. Thus the bases for finances need to address complexity and quality for different types of evaluations. Third, the different sizes of organizations and mandates, with implications for the types of evaluations conducted, suggest that one size does not fit all, and associated with this is the need to consider economies of scale.[46] This principle predisposes small organizations to spend more in relative terms than larger organizations. In this regard, the general view is that the range of funding between 0.5% and 3% of organizational expenditure is worth consideration depending on the mandate of the organization, the size of the organization, types of evaluations, and role of the function in institutionalization and support to strengthening decentralized evaluation, national capacities for evaluation, and partnerships in evaluation.

78. In addition to these factors, there is a need to think in strategic terms when costing the central evaluation function taking into consideration the complementarities and synergies with other functions that provide an evidence-base for organizational decision-making. Thus there is a need to consider the decentralized evaluation function, the work by oversight offices on performance assessments, and the role of knowledge-management in pulling all existing knowledge sets to develop lessons for the organization. All of these functions provide a diverse evidence base for organizational decision-making. It is also necessary to take into consideration the nature of each agency's operations: single location versus multi-locations, variability of operational contexts and service delivery versus administration and management.

79. This set of factors should be considered in the development of a normative framework for resource allocation. It is suggested that a typology of the work undertaken by the United Nations needs to be developed that allows specific evaluation approaches and related resources to be matched up. Selected members of UNEG see the need to take up the analysis and development of parameters and a benchmarking framework that would guide the development of budget norms for the allocation of resources for different types of organizations. That would be an important step that would benefit not only the United Nations system but also a

---

[46] Definition of economies of scale: "The cost advantage that arises with increased output of a product. Economies of scale arise because of the inverse relationship between the quantity produced and per-unit fixed costs; i.e. the greater the quantity of a good produced, the lower the per-unit fixed cost because these costs are shared over a larger number of goods. Economies of scale may also reduce variable costs per unit because of operational efficiencies and synergies. Economies of scale can be classified into two main types: Internal – arising from within the company; and External – arising from extraneous factors such as industry size." (from Investopedia.com)

large number of other development partners, including bilateral and multilateral organizations. Evidence on budget allocations and the applicable norms indicate that these development partners face the same challenges as those outlined above.[47]

Costing and budget structures

80. Given current challenges in defining budget norms for the evaluation function, the need to develop and monitor over time a fully costed programme of work as a basis for setting norms is noted as an important consideration, as required in the recent policies of WHO and UN-Women, which allocate 3 per cent of programme budget. What would, however, also be helpful would be to enhance the comparability of budget frameworks. Currently, budgets across organizations are structured differently and budget items with same word mean different things. For the evaluation function across organizations, **there is a need for a standard way of assessing the nature of its work, and then its relevant evaluation approach and attendant structural and resource requirements.**

Options to enhance the efficient use of funds

81. While more funding is needed to address the increasing evaluation demands, evaluation heads have indicated ways of collaborating and doing work together among organizations to enhance efficiency. These include pooling funds and using one provider to conduct certain common activities, a good example being the the quality assessment of evaluation reports. Lessons learned from the experience of the Department for International Development evaluation function in the United Kingdom, where one provider is used for all units, could be useful. In addition, lessons learned from the 44 joint evaluations done in the past five years by the United Nations system organizations could also provide a good basis for assessing constraints and defining cross-organizational collaboration and ways of enhancing efficiency on a system-wide basis. There are also lessons to be learned from the Active Learning Network for Accountability and Performance in Humanitarian Action initiative where, for example, the quality assessment of evaluation reports of its members is done by the network. In this regard, a stronger role of the UNEG in quality assessment could be envisaged. **It is highly recommended that evaluation units use lessons learned to enhance alignment, cooperation and collaboration across organizations in order to enhance the efficient use of resources and partnerships for evaluation.**

Resourcing small and underfunded organizations and the effective management of resources

82. The study shows that most resource increases for evaluation, although not commensurate with demands on the function, have occurred in already well-established functions. There is a need to focus on strengthening those offices that

---

[47] See OECD, "Better Aid: Evaluation in Development Agencies" (2010).

have lower levels of maturity, that are seriously underfunded and where staffing is not at high professional grade levels.

83. The pattern derived from the data analysis shows that the low performing organizations in Cluster 3 show a large variability in levels of funding of the evaluation function. The two modes of funding show an average of 0.07 for 4 organizations and 0.6 of organizational expenditure for another 4 organizations. This group in Cluster 3 has proportionately more staff hired at the lower level grades (P2-P3) (Please see tables in supporting materials.). The group also tend to use its pool of existing staff to carry out a variety of non-evaluation activities (See Table 8 below).

84. The groups also have a considerable number of small and co-located functions. The evidence is strong in showing that organizations where the function is co-located dedicate less resources than those where the functions is a stand-alone function. The evidence suggests the need to seriously consider increasing the financing of the evaluation function of co-located organizations. It also suggests that such functions, particularly where the evaluation function is fully integrated with other oversight or management functions, should enhance the effective use of the time of professional evaluation staff and to focus that time on doing evaluation as opposed to doing other tasks.

Table 6: Expenditure for different types of organizations and dedicated time to the core function of report production versus other tasks

Type of organization	**% of organizational expenditure on evaluation**	**% time on evaluation report production and quality enhancement *(relative to time spent on other activities of the unit)***
A. Levels of performance		
Cluster I ***High level of development***	**0.33**	**69**
Cluster II ***Average level of development***	**0.38**	**84**
Cluster III ***Below-average level of development***	**0.30**	**51**
B. Size of organization[48]		
Large organization ***(between US$ 1-5 billion)***	**0.18**	**68**
Medium-sized organization ***(between US$ 1 billion and 300 million)***	**0.17**	**73**
Small organization ***(less than 300 million)***	**0.33**	**53**
C. Structural location		
Stand-alone	**0.33**	**73**
Co-located within oversight	**0.14**	**55**
Co-located within management	**0.28**	**64**

---

[48] Note that the larger percentage of resources for small organizations is associated to a large degree with the concept provided above on economies of scale.

85. Taking into consideration the value of evaluation for organizational and development effectiveness and the risks involved in not carrying out evaluation, particularly in the current financial context, the implementation of the recommendations below would enhance the effectiveness, sustainability, added value, and efficiency of the function.

Recommendation 3

The legislative bodies should request the executive heads of United Nations system organizations to develop comprehensive budget frameworks and resource allocation plans for their respective evaluation functions, based on the cost of maintaining an effective and sustainable evaluation function that adds value to the organization. The plans should be submitted for consideration to the legislative bodies within existing budgetary and reporting mechanisms and processes.

Component 2: Independence with inclusion for enhanced credibility

86. Independence helps ensure the impartiality and objectivity of evaluation and thus engenders credibility. The results of the study show *that this is the most developed component among all the components of the function*, although it is still operating within the average level of maturity according to the standards established in this study. Table 7 below highlights the areas of strengths and weaknesses for the component. Based on analysis of findings, the study makes the following conclusions.

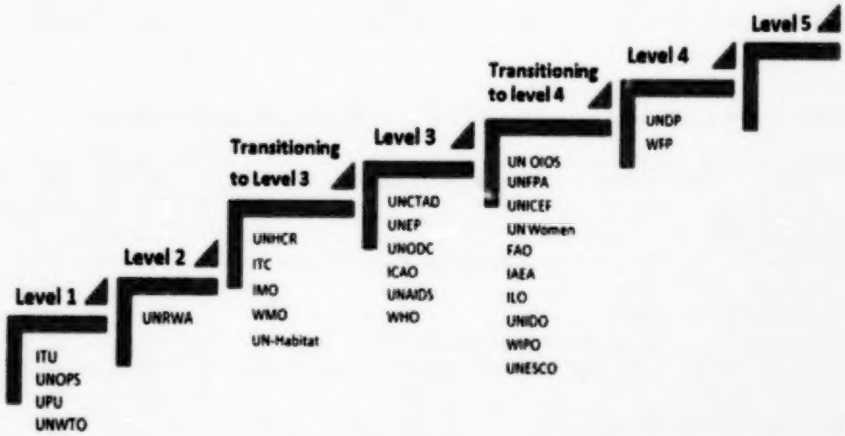
Table 7: Strengths and weaknesses for component on independence

Areas where progress has been made	Areas in which progress is lagging
Professional/technical independence*:** • Professional integrity and technical quality of the evaluators and managers of evaluation ***Structural independence: • Location of the evaluation office outside the office of the executive head and management and has independent decision-making capabilities ***Functional independence:*** • Head has full discretion over programme of work and in the development and issuance of report to member States and to the public and not through the Head of Organization • Organization and management provides full access to information ***Built in mechanisms for impartiality and transparency:*** • Extensive involvement of stakeholders for balanced perspective and transparency and use of independent external experts	***Structural independence*** • Appointment of head • Term of appointment of head and rotation in organization ***Functional independence*** • Independence of budgetary process

Conclusion 8: Despite advances made in system development to enhance the independence of the function, perceived credibility remains an issue with both governing bodies and individual member States. In particular, the system needs to address limitations associated with structural independence and in particular the appointment, tenure, diversity and professional education background among heads of evaluation units.

87. The independence of the function has been of interest to all stakeholders. The focus thereon has been driven by demands of member States for whom accountability, with independence, is critical for credibility. Evaluation units have also strongly focused on this area both from a professional point of view but also in connection with the interest to increase the use of their evaluations by bilateral organizations thereby limiting parallel evaluations by the latter.[49] Management has also played a most significant role in safeguarding independence. Twelve organizations have above average ratings on their level of independence. Six organizations are ranked at "average level" in ensuring systems are in place and operational. The two organizations with the most comprehensive systems for addressing all five criteria of independence - structural, functional, technical, transparency and behavioural independence are UNDP and the World Food Programme (WFP).

Figure IV. Level of maturity of participating organizations on independence of the function


Level 1
ITU
UNOPS
UPU
UNWTO
Level 2
UNRWA
Transitioning to Level 3
UNHCR
ITC
IMO
WMO
UN-Habitat
Level 3
UNCTAD
UNEP
UNODC
ICAO
UNAIDS
WHO
Transitioning to level 4
UN OIOS
UNFPA
UNICEF
UN Women
FAO
IAEA
ILO
UNIDO
WIPO
UNESCO
Level 4
UNDP
WFP
Level 5

88. One critical area of interest to many is the independence of the *reporting line*. The study encountered many challenges in defining what this means, as there are various interpretations and definitions of what this means for the independence of the evaluation function. In terms of direct reporting to governing bodies, only UNDP, UNFPA and OIOS have a direct reporting line. The Food and Agriculture Organization of the United Nations (FAO) also has a direct reporting line to both

---

[49] Ian C. Davies and Julia Brümmer, "Lessons-Learned Study of Peer Review of UNEG Evaluation Functions. Final Report", May 2013.

its governing body and the head of organization. The following example (outlined in tables 8 and 9 below) is, however, indicative of the interpretation of reporting lines, with its consequent effect on independence. For all organizations, the head of the evaluation unit is appointed by the executive head of the organization, with input in some cases from governing bodies. (See Table 9) Furthermore, 11 organizations have functions that operate outside the office of the executive head, independent of all decision-making processes. There are eight organizations with functions operating within the office of the executive head: they are not subject to line management decisions. On the independence of *evaluation reports*: 11 organizations issue *evaluation reports* directly to member States. Nine issue the *annual report on evaluation* directly to member States. Interviewees indicated that, where the report has to go through the head of the organization, there is no tampering with the evidence. At issue, however, is that, even where independence is not jeopardized and where current heads of organizations safeguard independence, the dependence of the function on management (subject to changes in leadership), may not be a stable solution for central corporate-level reporting.

Table 8: Reporting lines of the evaluation function of 24 organizations

Independence of reporting line	UNDP	UNFPA	UN OIOS	WFP	IAEA*	ILO	UNESCO	UNIDO	UNICEF	WHO	WIPO	WMO	UNCTAD	UN Women	FAO	ICAO	UNEP	UNODC	UNRWA	UN-Habitat	UNHCR	ITC	IMO	UNAIDS
1- Unit outside office of executive head, total indepence from all/almost all management functions	▨	▨	▨	▨	▨	▨	▨	■	■	▨	▨	▨	■	■	▨	■	■	■	■	■	■	■	■	■
OR																								
2-Unit within office of executive head	■	■	■	■	■	■	■	▨	▨	■	■	■	▨	▨	■	▨	▨	▨	▨	▨	▨	▨	▨	▨
AND																								
3- Evaluation report issued directly to member States without management involvement (sometimes via an oversight and advisory committee of the Board)	▨	▨	■	▨	▨	▨	▨	▨	■	▨	▨	▨	■	■	▨	■	■	■	■	■	■	■	■	■
4- Annual report on eval. issued directly to member States with no management interference and report is comprehensive	▨	▨	▨	▨	■	▨	▨	▨	▨	■	■	■	▨	▨	■	■	■	■	■	■	■	■	■	■
5- Program of work decided up and issued directly to member States (and getting approval from member States)	▨	▨	▨	▨	▨	▨	▨	■	■	▨	x	■	▨	■	■	■	■	■	■	■	■	■	■	■
6- Budgetary requests made directly to member States with no management interference	▨	▨	▨	■	■	■	■	▨	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
7- Administrative independence (including full control over management of financial and human resources, selection of staff and consultants, and approval of recurrent administrative functions	▨	▨	▨	▨	▨	▨	▨	▨	▨	■	■	▨	■	▨	■	▨	■	■	■	■	■	■	■	■

Mechanism is applicable	▨
Other mechanisms are applicable	■
No approval by member States	x

*IAEA manifests certain variations specific to the organization

89. One main concern is when a focussed effort on independence results in the central unit being totally isolated and removed from the organization that it is expected to serve. Independence alongside the engagement or inclusion of stakeholders rather than isolation is increasingly seen as critical for the credibility, trust and transparency of the United Nations system evaluation function. A good number of the organizations perform well in terms of this indicator. Mechanisms used to involve stakeholders in the various phases of the evaluation process provide important alternatives to attaching the function to the executive head if the purpose for so doing is to enhance linkage with the organization.

90. The major challenge for the function from the evidence compiled in this study is structural independence: the main issue requiring attention relates to the head of the evaluation unit and reporting lines. At issue is: (a) who selects or decides on the appointment of the head of evaluation; (b) the term of the head and whether he/she can rotate back into programme management after heading the evaluation function or while in office; (c) the background and education of the head on evaluation; and (d) regional and gender diversity, given current imbalances. Table 9 below presents the current status as reported by the various organizations.

91. These issues raised about the leadership of the evaluation function and its direct role in enhancing credibility and trust among a diverse range of stakeholders are challenging ones for most organizations. In the case of the selection and appointment of the head, neither selection by executive heads or by member States is free of political bias. In supporting the independence of the oversight function, JIU had made the recommendation that with respect to the **appointment of the head of internal oversight, the legislative bodies in each organization should decide that: (a) qualified candidates should be identified on the basis of a vacancy announcement that should be widely publicized; (b) Appointment should be subject to consultation and prior consent of the governing body; (c) Termination should be for just cause, and should be subject to the review and consent of the governing body; (d) A non-renewable tenure of five to seven years should be established, with no expectation of any further employment within the same United Nations organization at the end of the term.**[50]

92. The study examined only the last point and other factors that safeguard the independence of the head, such as reporting lines. The analysis shows a very mixed pattern of how this issue is addressed, reflecting a need for all organizations to re-examine the total set of indicators of independence and develop a rational and balanced approach in seeking to enhance and safeguard independence.

93. There are 13 organizations whose evaluation functions are not structurally independent of senior management and operate within the office of the head of organization (either as a stand-alone evaluation function or fully integrated into the management in the office). Six of these have no independence in reporting (UNEP, UNODC, the United Nations Relief and Works Agency for Palestine Refugees in the Near East (UNRWA), the United Nations Human Settlements

---

[50] Oversight Lacunae in the United Nations system, JIU/REP/2006/2.

Programme (UN-Habitat), the International Trade Centre (ITC) and the International Maritime Organization (IMO). Five have limited independence in reporting (the International Civil Aviation Organization (ICAO), UN-Women, the United Nations Conference on Trade and Development (UNCTAD), the United Nations Children's Fund (UNICEF) and UNIDO), with variations as observed in table 8 above. 11 of the 13 of the heads of evaluation have fixed-term appointments and rotate within the organization after leaving the function. Two have term limits: UNICEF, where the head has two terms that cannot be renewed after the second term, and ICAO, where the head has one term of between five and seven years, and both of whom have limited independence in reporting lines. **The legislative bodies of these organizations should re-examine their policies for the structural independence of the evaluation function and decide on how best to enhance and safeguard structural independence, including in the appointment of the head of the evaluation unit.**

94. Faring much better on structural independence in terms of: (i) the functions being located outside the executive head office;(ii) the appointment of the head respecting independence; and (iii) independence in reporting, are nine organizations: UN-OIOS, UNDP, UNFPA, WFP and UNESCO, UNDP, the World Intellectual Property Organization (WIPO), and the Inspection and Evaluation Division of OIOS. Of this group, however, OIOS. WIPO and UNESCO have the strongest independence of head of evaluation with one term of five and six years, respectively, non-renewable with no re-entry into the organization. A one-term, non-renewable policy is considered to grant more independence than a renewable one. Renewal in the majority of cases is granted by the head of the organizations and this has its effects on limiting independence when the incumbent seeks to remain in the organization. A five to seven year period provides an opportunity to operate over time with full capacities in understanding the organization and being able to best position the function to support change and development. From a human resource management perspective and for applying an appropriate incentive to attract best talent from the outside, a minimum of 5 years of service is needed to benefit from the UN pension fund. **It is recommended that legislative bodies of UNDP, UNFPA and WFP re-examine the policies of two non-renewable terms and decide to put into effect one non-renewable term of five to seven years.**

95. Organizations not included in the lists in the two preceding paragraphs include ILO, WHO, the World Meteorological Organization (WMO), FAO, the Office of the United Nations High Commissioner for Refugees (UNHCR), the Joint United Nations Programme on HIV/AIDS (UNAIDS) — most of them are specialized agencies with varying structural independence, leaving one to question whether the choices made are based on rational reflection or tied in some way to mandate or size. ILO has an office outside the office of the executive head, but the head of the evaluation unit is typically a longer-serving staff member and can rotate within the organization. The FAO head of evaluation can complete two non-renewable terms, but has practically no independence in reporting. The head of evaluation of WHO operates outside the office of the executive head but the appointment allows rotation within the organization. The head of WMO

evaluation operates outside the office of the head of the organization, but with severe limitations on the independence in his reporting lines. The heads of evaluation of UNHCR, UNAIDS, ITC and UNCTAD are fully integrated into management units, have no independence in their reporting lines and are employed on a fixed-term basis. **It is recommended that the legislative bodies of these eight organizations re-examine structural independence, the hiring of the head of evaluation and rotation in the organization to enhance and safeguard the independence of the function with due regard to organizational conditions and experiences of comparable organizations.**

Table 9: Structural independence: how organizations are distributed based on appointment regime and rotation in the organizations

	Appointment by governing/ legislative bodies	Appointment by executive head with input from governing body and/or other independent human rights or selection mechanism	Appointment by head with no input from governing bodies or other parties
Term of office head has one term of between five and seven years and cannot return to the organization		Secretariat of the United Nations (5) WIPO (5) UNESCO (6) ICAO (5)	
Head has not more than two terms and cannot return to the organization		UNDP (4+4) UNICEF (4+4) WFP (4+4) FAO (5+5) UNFPA (5+5)	
Fixed-term appointment and same rules apply as rest of organization for rotation within the organization		UNODC WMO UN-Women WHO**	UNIDO, UNCTAD, UNEP, UN-Habitat, UNHCR, UNRWA, ILO, IMO, UNAIDS, IAEA*

*IAEA policy for all staff is for one term of seven years and exit the organization, but rotation within the organization is allowed with effects on independence.

**WHO is undergoing restructuring of its evaluation function in 2014 and new policies are being developed.

96. With the increasing complexity of evaluation and given better understanding of it as a profession and discipline with defined standards and rigour in applying complex methods, questions are raised about the education and background of heads as evaluators. Heads at D1, D2 and P5 levels come from diverse backgrounds. Heads of the evaluation functions within oversight units,[51] who are also heads of oversight and make up about 38 per cent of heads in the 24 JIU participating organizations, are auditors. Most, but not all, heads of other evaluations are appointed based on an evaluation background. In some cases, solid

---

[51] There is in 2014 one exception: the new head of the oversight office of UNESCO is a professional evaluator.

institutional knowledge has been seen as an important edge over evaluation knowledge. It has been stated that the evaluation background of heads is not as significant as having the head supported by evaluators. At the same time, leadership and credibility may be hampered by the absence of the critical and strategic mode of inquiry.

97. Limiting the credibility and status of the evaluation function, in the current context of the United Nations with rigid hierarchies, is when the evaluation head is assigned a P5 level To remedy this situation there is now an increasing trend in larger organizations to hire heads of evaluation at D1 level and above who have an education or solid experience in evaluation in addition to other leadership requirements. Finally, the analysis shows that demographic information about heads now raises questions about the degree to which the selection of heads takes into account the United Nations principles of inclusion and diversity. Eighty-five per cent of heads are from developed countries, and 75 per cent are male.

98. The implementation of the recommendation below will enhance the credibility and effectiveness of the evaluation function. Given the wide variations, described above, in the configuration of independence among the various organizations, the recommendation, while generic, has to be implemented with due regard to the various types of organizations.

Recommendation 4

The legislative bodies should direct the executive heads of United Nations system organizations to review and revise, as necessary, existing policies for the appointment of the heads of evaluation offices, in order to enhance independence, integrity, ethics, credibility and inclusion, with due regard to the following criteria:

- **Term limits should be established for a single non-renewable term of office of between five and seven years, with no possibility for the incumbent of re-entry into the organization;**
- **The head of evaluation should have qualifications and substantial experience in evaluation, complemented by experience in the related fields of strategic planning, basic and operational research and knowledge management, and should have excellent management and leadership attributes.**

Component 3: Quality — technical and managerial rigour for enhanced credibility

99. Table 10 below provides a summary of the strengths and weaknesses of the indicators assessed for this component as defined in the maturity matrix.

Table 10: Quality and credibility of evaluation processes and products — strengths and challenges

Areas where progress has been made	Areas in which progress is lagging
▪ Internal quality assurance of evaluations (internal peer group, involvement of reference group from programme units; involvement of advisory group for wider perspective and global value) ▪ Staff competencies: technical, professional, managerial and institutional knowledge	▪ External independent, quality assessment of reports and of the function (e.g. UNEG/OECD DAC peer review and other independent reviews requested by boards) ▪ Quality of evaluation reports ▪ External assessment of the function: UNEG peer review for small and co-located organizations ▪ Professional development of staff ▪ Innovations with methods and types of evaluation ▪ Technical and managerial guidelines and tools

Conclusion 9: Organizations have sought to ensure validity and rigour in the evaluation function through a range of measures, including internal unit quality assurance, the use of external experts, the use of a reference group of key stakeholders and enhanced staff competencies. While organizations use a variety of internal quality assurance mechanisms, missing is the use of two external measures which are significantly related to a high level of development of the evaluation function. These are (a) the external assessment of the quality of the function, including via the UNEG/OECD DAC peer review of the function and (b) the independent assessment of the quality of central evaluation reports. There is limited independent evidence available on the quality of evaluations produced by the central evaluation functions of the United Nations system. This limits the credibility of the function.

100. The study found that the evaluation functions of a large number of organizations apply an extensive number of measures to enhance validity and reliability. Internal unit quality assurance mechanisms within evaluation offices include: using peers in the unit or other organizations to be part of a quality assurance team for an evaluation; using external advisory bodies; using reference groups of key stakeholders; and holding public and stakeholder meetings at various stages of the development of the report to validate the findings and conclusions. These all contribute to enhancing the quality and credibility, as well as the utility, of the evaluation reports.

101. Next in line in bolstering the credibility of the evaluation function are the competencies of staff. Their technical, professional, managerial, institutional and programmatic knowledge and their high standards of professional and personal integrity are all reported to ensure independence and enhance balanced and objective evaluations. Core staff of the evaluation unit generally plays a leadership role in ensuring application of the evaluation principles for technical rigour and in ensuring that the evaluation is meaningful to the organization, while also managing the effective processing of the evaluation. Staff carries out evaluations together with contracted external experts/consultants who enhance the content

validity of the evaluations in cases where the unit covers a broad range of subject matter that cannot reasonably be housed on a permanent basis.

102. While the internal (unit) quality assurance mechanisms described in paragraph 100 above are quite robust, the system lacks credible and transparent, independent, external quality assessment of the evaluation function and its evaluation reports based on well-defined, objective and systematic methods. Nine organizations have had an external assessment of their functions using either UNEG peer review mechanisms and/or external evaluations commissioned by the executive board or management, or by bilateral groups.

103. For the assessment of the quality of evaluation reports, the study did not include an actual meta-evaluation of the reports. As for most other rubrics in the study, the Inspectors analysed quality based on existing external assessment of the quality of evaluation reports. Thirteen organizations have undergone external assessments of the quality of their reports. Such assessments indicate that only seven of these organizations have reached a high level of quality of evaluation reports meeting professional standards (ILO, OIOS, UNIDO, UNICEF, UNDP, UNEP and WFP). The various organizations do draw upon the UNEG standards for quality reports, but they also apply other criteria and use different measurement systems. This makes interpretation and comparative analysis difficult.

104. The limited amount of systematic external assessment and knowledge about the quality of evaluation reports, and the less-than-desirable quality of evaluation reports, raise some major concerns about the outputs and value of the function of the system as a whole. As in many other areas of assessment, however, there are variations among the organizations. There is an increase in the systematic external assessment of the evaluation function and its products by bilateral organizations. While this is important, a proliferation of different external assessment systems, not all using the same criteria or standards, could result in a duplication of effort, high transaction costs and work at cross-purposes.

105. The implementation of the recommendation below will enhance the credibility and hence effectiveness of the function.

Recommendation 5

The Secretary-General of the United Nations, in his capacity as chair of the United Nations System Chief Executives Board for Coordination (CEB), should request UNEG to collaborate in developing a robust and harmonized quality-assurance system for the evaluation function across the United Nations system.

Component 4: Utility — use and impact of use

106. The study assessed the following areas that theoretically are critical for enhancing the outcomes and impact of the evaluation function: (a) the conditions in place to enhance use, including tools, techniques and processes; (b) quality

reports; (c) stakeholder involvement; (d) organizational culture and leadership; (e) the level and nature of use of evaluation products; (f) the complementary roles and responsibilities of evaluators, managers and governing bodies in enhancing use; and (g) the systems in place for assessing the impact of the use of evaluation. The organizational culture and leadership, both of which are most significant for the use of evaluation, were addressed in an earlier section and are thus not covered in this section.

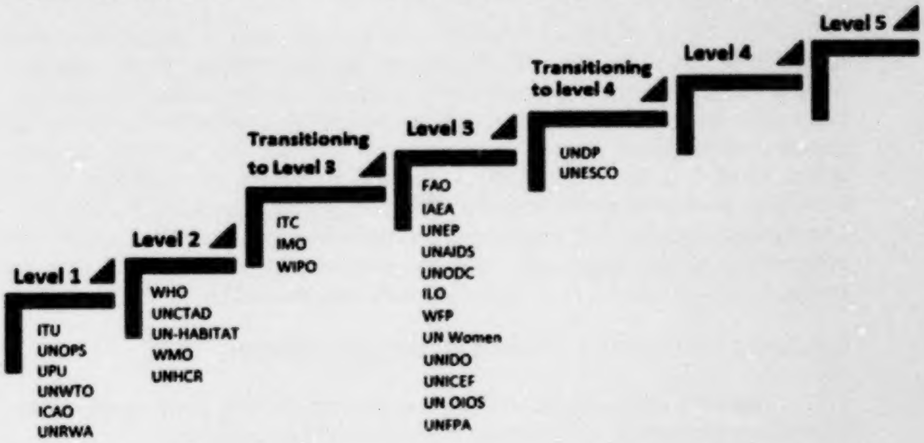
107. The results of the assessment show that the component on the utility of the function is the weakest of all five components of the function, operating below average level in the aggregate (the exceptions being those indicators listed in the left column of table 11 below which provides indicators for which progress is observable).

Table 11: Strengths and weaknesses related to utility of the function

Indicators where progress has been made	**Indicators in which progress is lagging**
Measures that enhance use ▪ Demand and intention to use and planning criteria that are relevant and respond to demand ▪ Stakeholder involvement Targeted tools, systems and measures in place for direct effect ▪ Timeliness and feeding into decision-making processes ▪ Accessibility of reports in websites and other platforms for use and transparency ▪ Sharing internally	Use of evaluation (outcome indicator) ▪ Level of use via recommendation implementation rates Influence of the prevailing culture ▪ Culture of evaluation (organizational culture; leadership; alignments in knowledge system) Targeted tools, systems and measures in place to enhance use (direct effect) ▪ Dissemination and communication strategy ▪ Recommendation tracking system (for follow-up implementation of use of evaluation) ▪ Sharing externally Impact assessment system: ▪ Organizational visions and strategy and how evaluation supports this ▪ Impact indicators for evaluation function ▪ Impact follow-up system and reporting

108. The scores for organizations on this utility component show that there are no organizations at the high end of the scale. Two organizations are at an “above average” level, 15 organizations at an “average” level and the rest (11) at a “below average” level.

Figure V. Performance of 28 JIU participating organizations on utility


Level 1
ITU
UNOPS
UPU
UNWTO
ICAO
UNRWA
Level 2
WHO
UNCTAD
UN-HABITAT
WMO
UNHCR
Transitioning to Level 3
ITC
IMO
WIPO
Level 3
FAO
IAEA
UNEP
UNAIDS
UNODC
ILO
WFP
UN Women
UNIDO
UNICEF
UN OIOS
UNFPA
Transitioning to level 4
UNDP
UNESCO
Level 4
Level 5

Conclusion 10: The systematic use of evaluative evidence for decision-making is low. It does not match the level of effort that has been directed at establishing the evaluation mechanisms and systems to enhance the use of evaluation by the central evaluation units. Limited use of evaluation does not allow for enhanced, evidenced-based decision-making and improvement of practice. It thus limits the value of the function. Organizations have not focused on the impact of the use of evaluation. Systems for recording and assessing impact are rudimentary where they exist. The conceptualization of the impact of the use of evaluation would require the development of a comprehensive and integrated vision and strategic framework for the function, its expected results and indicators of impact, and how this impact would be aligned with results of other oversight and knowledge-production systems in the organization.

Level of use and follow-up tracking systems to assess the use of evaluation

109. A concrete way of assessing the use of evaluation is by examining the use made of the recommendations from evaluations. The level of use of the recommendations from evaluations is well below the average level of expected performance. Only 21 per cent of organizations have a high level of use. This means that more than 85 per cent of recommendations are implemented within a three-year period. This finding is based on reported information derived from the implementation rates in follow-up tracking systems. The tracking system for monitoring the implementation of recommendations provides a concrete basis for assessing the level of use of evaluation findings and the recommendations or issues raised in evaluations. Forty per cent of organizations (11)[52] have well-established tracking systems with good reporting on the status of the use of evaluation.

---

[52] OIOS, UNDP, UN-Women, UNICEF, IAEA, ICAO, UNESCO, UNIDO, ILO, WIPO and WMO.

Nature of use

110. Five types of use (focused on instrumental use) of evaluations were identified in this study. Use; (a) for strategic decision-making; (b) for corporate summative (programmatic or institutional) purposes; (c) for formative purposes – programme and project design and adjustments; (d) by external partners (national partners, development partners or professional groups); and (e) use to enhance added value (e.g. via meta-analyses and lessons notes), to contribute to the evaluation profession and to advance methodology and approaches in evaluation. The findings indicate that evaluations are most frequently used to improve the programmes of the organization and the achievement of results and then for strategic decision-making (see supporting materials, annex 21).[53]

Conditions put in place to enhance the use of evaluation

111. The study examined the conditions and elements put in place by the various organizations to ensure a high-impact evaluation system, including:

- the influence of the prevailing culture and leadership;
- the demand and intention to use the system and its coverage;
- a well-defined and direct set of critical tools and processes to enhance use and impact.

112. The analysis shows that the most developed mechanisms of the evaluation system to enhance use are: efforts to make coverage relevant and meet demands; the involvement of stakeholders; accessibility and transparency via websites and evaluation databases; and the timeliness of evaluation reports. In general, systems and conditions in place for enhancing use of evaluation are at an "average" level of development and are being institutionalized. However, more needs to be done particularly by management to enhance the culture of evaluation and learning. ILO has initiated a focused and systematic process of enhancing use and this needs to be documented and lessons from it shared on a wider scale.[54]

Systems in place for assessment of the *impact of evaluations are at a rudimentary level*

113. It was realized during the scoping for the study that, while anecdotal evidence exists, the assessment of the impact from the use of evaluation and its products by the various organizations was almost non-existent. The study thus assessed the systems in place for assessing impact. The evidence shows that the following are missing when systems are developed: (a) a vision and strategy for the evaluation function that includes a results framework and a comprehensive set of indicators for the evaluation function, and the alignment of the function's results framework with organizational results; (b) mandatory and well-established

---

[53] See www.unjiu.org/en/Pages/default.aspx.

[54] Janet Neubecker, Matthew Ripley and Craig Russon, "Building for Utilization: The Case of the International Labour Organization (ILO)" in Marlène Läubli Loud and John Mayne (eds.), *Enhancing Evaluation Use: Insights from Internal Evaluation Units* (Sage, London, 2014).

implementation of recommendations and a follow-up tracking system; (c) reporting on implementation status and on the effects as per the established results framework; and (d) reporting on other factors that are also contributing to the effects and impact of use. Organizations need to make the use of evaluation to influence accountability, learning and decision-making, and the impact of the use of evaluations, a strategic priority. Related is the recommendation on the need to develop a learning organization that uses evaluative information for change and development of the organization and its ability to deliver better results. Organizations should also report on the impact of the use of evaluation in the annual reports on evaluation submitted to the legislative bodies.

114. The implementation of the recommendation below will enhance the effectiveness of the evaluation function.

Recommendation 6

The executive heads of United Nations system organizations should make the use of evaluation an organizational priority based on a well-defined vision, strategy and results framework for the evaluation function, and report to their legislative bodies on the level, nature and impact of use of evaluation.

Component 5: Relevance and readiness to support United Nations organizations and system-wide reforms and to address emerging changes and challenges

115. This component addresses (a) relevance and efficiency via joint evaluation and (b) the capacity and strategic positioning of the evaluation function in helping the United Nations system to address changes, challenges and emerging agendas for transformative change and sustainable development, both within organizations and at a system-wide level.

116. In the current context of change and challenges in the world, questions have been raised as to the relevance of the evaluation function to support the United Nations system at the crossroads of development.[55] One issue raised is the disaggregated and functional structure of the United Nations system, which would limit the move towards a system-wide perspective that would support emerging agendas, including the post-2015 development agenda for sustainable development. It equally limits the development of a system-wide perspective on evaluation that connects with and influences the evaluative work carried out in the individual organizations. In other words, individual corporate central evaluation functions, at a minimum, would be "sub-systems" of a "United Nations-wide system". Yet the reality is far from this. In such a scenario, one would expect to see linkages among central evaluation functions that influenced their individual work programmes and ask questions such as - what type of evaluation would add the most value from the perspective of the United Nations as a whole, rather than the individual organization, and which in turn influenced decisions on resource allocations between the various central evaluation functions.

---

[55] Bruce Jenks and Bruce Jones, "United Nations Development at a Crossroads".

117. Yet the debate on the post-2015 development agenda to date is conspicuously silent on the need for evaluation that meets such criteria. A focus on data and data reporting as part of an accountability framework for reporting overshadows an intense reflection on the implications of the sustainable development agenda for the role of evaluation, including the critical role of decentralized evaluation in supporting formative evaluation and mid-course correction. It equally sidesteps the importance of national systems in evaluation and evaluation-capacity development, and the role of the United Nations system in strengthening such capacity.

118. The study examined and assessed how the various organizations are addressing relevance to organizational demands for coverage, including through joint evaluation, while also addressing a number of factors that are significant for responsiveness and readiness in supporting reforms and changes in our time.

Conclusion 11: The evaluation function across the 24 organizations reviewed has sought to coordinate and harmonize work mainly through UNEG and to apply collectively developed norms and guidance primarily within individual organizations. It has yet to collaborate and conduct joint work across agencies. The "*disaggregated and functional*" nature of the United Nations system with separate governance, operational procedures and budgeting processes hinder efforts for joint evaluation. The evidence suggests that *joint programming* provides an opportunity for joint evaluation, typically in cross-cutting themes such as food security, gender, the environment and labour issues.

119. The Inspectors found that about 50 per cent of the organizations of the United Nations system are not involved in joint evaluations; most do not consider it a priority and have focused more on placing the evaluation function on the right track. The perceived difficulties of joint or inter-agency work also limit motivation for joint work. Eleven of the 24 organizations have carried out joint evaluations: a total of 44 joint evaluations have been carried out, principally among these United Nations agencies over the past five years. One normative agency, ILO, carries out advisory, coordination and support functions for on-going evaluations of other agencies, reflecting a different form of inter-agency work on evaluation for predominantly normative organizations. ILO has conducted over 150 joint activities of this type over the past five years.

120. Interviews with groups of organizations involved in joint evaluation highlight a continued commitment to doing joint evaluations, but also a large number of structural and systemic issues which affect joint and inter-agency evaluation. These include: separate governance structures; the need for upfront investment costs for initiating joint evaluations (both financial and time, given the huge transaction costs); the need for support in administration and coordination to permit agencies to focus on substantive aspects of the joint work; difficulties in synchronizing programme and budget cycles; and differences in approaches between organizations; different legislative bodies that would receive and discuss the joint report; uncertainty as to where the joint evaluation work is taken up at level of organization or inter-agency body; and poor management buy-in and follow-up.

121. One significant factor that facilitates joint evaluation among central evaluation offices is joint programming for cross-cutting issues. A joint programme is a set of activities contained in a common work plan and related budget, involving two or more United Nations organizations and subnational partners.[56] This is the case for gender equality, food security and livelihood protection, and decent work, where joint evaluations have been conducted (see supporting materials, annex 26).[57] Lessons learned from such joint evaluations at corporate central level could play a significant role in advancing the agenda for coherence in evaluations.

122. Besides joint programmes, UNDAF itself as a coordination mechanism should also increase joint evaluations of country programmes, including evaluations done by the central function. The evidence suggests this has not yet been attained. Decentralized evaluations at country level do not generally feed into UNDAF. In a similar way, there is little or no coordination between (a) decentralized evaluation, (b) UNDAF evaluation, and (c) the strategic country programme evaluations done by the central evaluation units. The Inspectors could find only one case of an effort by UNDP to pilot a coordinated evaluation of this type. It is reported that managing this type of coordination is very difficult and that the success of such joint evaluations require the leadership role of the government. **It is recommended that executive heads of the organizations of the United Nations system fully support their evaluation offices actively engaging in joint evaluation.**

Conclusion 12: The findings indicate that the evaluation function is not as robust as it ought to be for effective strategic positioning of the evaluation function to support the United Nations system. The level of the six most mature central functions is assessed as "adequate" in meeting evolving demand. For the remaining 18 organizations, the challenge remains in their ability to grasp and use opportunities presented by the evolving global context to enhance collaboration, innovation and working across organizations and thus enhance their relevance.

123. The following are open-ended comments made by heads of evaluation units to questions on organizational reflections and plans for moving forward in the future: "…being strategic in reflection and outlook; focused on strategic direction-setting; being resilient in responding to change; playing a catalytic as opposed to a reactive role to enhance the relevance of the United Nations; being innovative and engaging in constant adaptation in approaches and methods and priorities, in refinement and renewal in response to a fast pace of change; engaging global platforms and enhancing partnerships including partnerships with analytic centres and global think tanks now driving the knowledge agenda; and having nimble ways of working together"

124. These attributes, as expressed by evaluation heads and staff across the system, are most important for the role the evaluation function could play in supporting the United Nations system as it seeks to carve out a comparative niche and a path forward in a context of global changes and demands for sustainable

---

[56] United Nations Development Group, Guidance Note on Joint Programming, December 2003.

[57] See www.unjiu.org/en/Pages/default.aspx.

development. These attributes were examined in the study via the various rubrics of the maturity matrix. Some of the rubrics are also the same as those that operationally define the learning organization given the close relationship that exists between the two. Table 12 below shows the key dimensions and indicators from the maturity matrix used for assessing the relevance, responsiveness and readiness of the evaluation function. The numbers in the table indicate the mean score across the organizations, which is indicative of level of readiness in supporting the United Nations system. Overall, the level of readiness or capacity for addressing changes and challenges is below average. On average, the evaluation functions of only six organizations are assessed to be ready to address the challenges. They are ILO, UNICEF, UNDP, UNIDO, UN-Women and WFP. They are assessed to have the requisite outward-looking focus and a definite intention of addressing global changes and emerging imperatives. There are, however, a good number of organizations that are operating at an average level of readiness, and this in spite of the focus they are giving to enhancing the coverage and quality of evaluations.

Table 12: Subcomponents and indicators used to assess relevance and readiness

COMPONENT 5: Capacities for enhancing relevance, responsiveness, adaptability and readiness for change and challenge (4.8): (Highest means score is 10)
Subcomponents
Strategic in reflection, approach and mode of operation (mean score 4.9)
Relevance in covering programmes and projects of the organization (evaluation planning and coverage): having a vision and strategy for the function anchored in the organization and the United Nations system as whole; leadership; member States' roles and understanding and appreciation of the added value of evaluation; strategic use of evaluations and impact of use
Innovation, adaptability and renewal/flexibility and resilience (mean score 5.3)
Adaptability and continuous improvement and growth — dynamic function: balancing a range of demands: continuous assessment and adaptation; contribution to advancing development evaluation and development of appropriate methodologies that are valid for the types of United Nations system interventions
Coherence, inclusion, ownership and alignments (mean score 4.4)
Responsiveness to support coherence, inclusion, ownership and sustainability with a focus on: (a) decentralized evaluation; (b) United Nations reform for coherence including doing joint evaluation; (c) gender and human rights; (d) national evaluation capacity development and alignments with national systems and enhancing ownership
Responding to global changes and challenges (mean score 4.7)
Addressing global changes and challenges in evaluation and engagement on global platforms and global governance issues, including sharing knowledge and directing change: development of the learning organization and culture of divergent thinking, innovation and risk-taking for organizational transformation; active in UNEG with a proactive role for coherence and as driving force for change not only in organizations but also in the United Nations system and globally

125. Notwithstanding the current level of readiness, managers of evaluation and of programmes express a very clear intention for the function to operate with alternative and more dynamic models:[58]

> "As a significant agent of change: participating in sector-wide and other evaluations of a more strategic nature; enhancing the institutionalization of decentralized evaluation for the complementary role it can play to add value; having common shared systems (common database of external consultants, common systems for quality assessment of reports, common risks frameworks); working together and limit silos in leaving each organization doing it alone; having direct peer exchange among evaluators; implementation of system-wide evaluation policy and coordination and joint planning."[59]

126. These ideas are indicative of the onset of a dynamic process. They echo the principles emerging from the debate on the post-2015 development agenda:[60] interdependence and interconnectedness; crossing boundaries in working together; doing things that are fit for purpose and adding value and enhancing linkages, including national capacities. They reflect a new mode of operation; one that is more flexible and yet requires collaboration and action among all parties.

127. Yet, a response to ensure that the evaluation function has these attributes cannot be ad hoc. It requires (a) a defined strategy for prioritization; (b) collaboration on various fronts, which require breaking restrictions when trying to work across organizations in driving transformative and interdependent change; and (c) the development of renewed incentives systems. Heads of evaluation and managers note that such a response goes beyond coordination and harmonization of activities under UNEG. It needs to be supported by adequate resources — financial and human — and intellectual space for divergent thoughts and exploration of alternatives, commitment and time need to be provided.

128. The independent system-wide evaluation of operational activities for development referred to in paragraph 14 above presents an opportunity for the organizations of the United Nations system to collaborate with JIU, which has a mandate for independent system-wide evaluation. A policy of independent system-wide evaluation provides a framework for partnership in system-wide evaluation of operational activities for development, building on existing systems, mechanisms and capacities, The principles for evaluation of the policy — subsidiarity, economy and efficiency — provide an explicit basis for coordination and collaboration across all levels of the system, including on the effective use of the pool of existing evaluations. The modalities for evaluation include (a) the use of the existing body of knowledge produced at central and decentralized levels to

---

[58] See content analysis of responses of heads of evaluation and other mangers to open-ended question on maturity matrix.

[59] Responses of heads of evaluation units and staff to open ended questions on reflections on the direction of the evaluation in moving forward into the future.

[60] The High Level Panel on the Post-2015 Development Agenda, "A New Global Partnership: Eradicate Poverty and Transform Economies through Sustainable Development" (United Nations, 2013).

carry out meta-synthesis at system-wide level; and (b) coordination and collaboration among various central units working under a common framework in carrying out evaluation around a theme for use at both the organizational and United Nations system-wide level.

129. A policy of independent system-wide evaluation, when implemented, has a role to play in enhancing collaboration and efficiency and in providing a more holistic picture of United Nations-system coherence and of the effectiveness of operational activities for development. Efforts to start the pilot of two evaluations have, however, been thwarted by the lack of funding despite the General Assembly resolution requesting extra budgetary support.[61] The success of the Scaling Up Nutrition initiative in engaging a broad range of partners, including from the private sector, might provide lessons for the financing of independent system-wide evaluations.

130. The implementation of the recommendation below will strengthen coherence and harmonization.

Recommendation 7

The executive heads of United Nations system organizations should request evaluation offices to reassess the policies, strategies and priorities of the evaluation function. They should strategically position the evaluation function in their respective organizations so as to enhance its relevance in enabling the United Nations system to address current changes and challenges, and to achieve impact and sustainability.

National evaluation capacity for sustainability of the evaluation function of the United Nations system

Conclusion 13: General Assembly resolutions on national evaluation capacity development go back to 1989. The organizations of the United Nation system recognize the value of such capacity development and view alignments with national systems for evaluation, including for data and statistics, as critical not only for countries but also for the sustainable development of the evaluation function of the United Nations system. The current level of engagement and effort in supporting national evaluation capacity development including enhancing ownership of evaluation findings is rudimentary. Many organizations do not see themselves as having a mandate for national evaluation capacity development. Others do not see it as a responsibility of the evaluation function but as a programmatic responsibility. Very few include national evaluation capacity development in their evaluation policies. Strategies applied by the various evaluation units give prominence to the use of national consultant services in the context of on-going central evaluations. The exceptions are three evaluation units (from UNDP,

---

[61] A/RES/68/229.

UNICEF and UN-Women) that have been given such a mandate in their evaluation policies and also have a vision and strategy or plan and carry out planned activities.[62] In summary, the current approach across the United Nations system does not operate from an integrated framework and has yet to build on the understanding of capacity development and national ownership as an endogenous process.

131. Enhanced national capacity for evaluation is not only of value to countries but it is also a *critical success factor of the evaluation function of the United Nations system.* This was the operating premise for examining how the evaluation function sees its alignment with efforts directed at strengthening national capacities for evaluation. The Inspectors are fully aware of the debate on the fact that national evaluation capacity development is partly a programmatic effort that should be tied to the project cycle and management, as well as to interventions directed at enhancing governance and public-sector management. The JIU report in 1982[63] examined the various programmatic efforts by 23 United Nations agencies and the International Bank for Reconstruction and Development, OECD and the European Commission, highlighting actions taken and challenges. A future JIU study would, in the current context, be a commendable effort in assessing advances made by the various organizations.

132. Programme units have acknowledged their own limited capacities and have looked up to the evaluation community for support in developing national evaluation capacity. Evaluators of the United Nations system recognize the strategic importance of national evaluation capacities for the impact and sustainability of the evaluation function. As per the views of a large number of the heads of evaluation units, alignment of the United Nations evaluation function with national systems and development of capacity is significant for many reasons: (a) mutual accountability for results; (b) inclusion of diverse knowledge systems to enhance validity; (c) involvement in planning and management to enhance use and impact; (d) engagement to enhance harmonization and mutual capacity development; and (e) engagement with existing systems and institutions on the ground to enhance the efficiency and sustainability of the evaluation functions of both the United Nations and countries, this being most critical in the current context. These views were expressed by both those who identify as having a mandate for national evaluation capacity as well as those who indicate they do not have a mandate but see national evaluation capacity development as important for the utility, impact and sustainability of the evaluation function of the United Nations system. As such, there was no objection to being assessed on their efforts to support national evaluation capacity development.

133. Analysis of the reported information indicates an increase in the linkage between the central evaluation function and national systems and national evaluation capacity development. At the current time, alignment with national

---

[62] UN-Women in 2014 also had a specific global evaluation strategy addressing national evaluation capacity and global outreach and partnerships.

[63] United Nations System Co-operation in Developing Evaluation by Governments, JIU/REP/82/12.

systems is not a consideration for 8 organizations. The other 16 organizations show different degrees of engagement with national experts and institutions. For some organizations, the activities carried out are typically not guided by a strategy, nor is their coordination among the various organizations. The primary mode of engagement is to include national consultants or institutions in evaluations being conducted. This is done in the majority of cases to enhance some level of ownership with effects on the use of evaluation but also to enhance capacity via action learning, engagement and participation. It is not an insignificant effort with effects on mutual capacity development. But more needs to be done in a more systematic and less ad hoc manner.

134. As stated above, there are two organizations, UNDP and UNICEF, which have made national evaluation capacity development an integral part of the operations of the central evaluation function. While the activities of these organizations for national capacity development are in general commendable, they are not developed according to a framework that would enhance collaboration and build on each other's comparative and added value in supporting national evaluation capacity development. Alignment of the work of the central function with the work of their programme units is also part of a strategy only now emerging. More recently, in 2014, UN-Women made national evaluation capacity development a strategic area of its global evaluation strategy. The global strategy is commendable in addressing broader partnerships with United Nations agencies, international financial institutions and private foundations.

135. What are some of the lessons to be learned on what makes for success in national evaluation capacity development? One most significant condition for success is the ability to place support for national evaluation capacity development in the context of country realities and to use home-grown knowledge and accountability systems.[64] Such an approach is critical in efforts directed at using the more recent guidelines on national evaluation capacity development provided by UNEG.[65] The UNDP model,[66] which builds on the endogenous capacity and leadership role of countries, is recognized as an exemplary approach to supporting national ownership in strengthening capacities. Evaluating the success of such approaches is full of challenges given that they are slow, amorphous, and require serious ethnographic studies of the development process and the types of regressions as well as quantum leaps that are made.

136. Other important consideration in enhancing national capacities include the use of non-traditional and private-sector mechanisms, such as venture capital schemes working with country and regional high-level experts to build up capacity. Equally important is working with existing national and regional knowledge-development institutions such as universities, analytic centres and

---

[64] Sukai Prom-Jackson, "Rebirth, Restoration, Reclamation, and Responsibilities of the Evaluation Function of Africa", in *Evaluation matter: How can we strengthen national evaluation systems – Perspectives from around the globe*, Quarterly Knowledge Publication of the African Development Bank, vol. 2, No. 3 (September 2013), pp. 76–81.

[65] UNEG, National Evaluation Capacity Development: Practical tips on how to strengthen National Evaluation Systems (2012).

[66] UNDP International Conferences on National Evaluation Capacities (2009 and 2013).

think tanks, including those being developed through the World Bank project on regional centres for learning on evaluation and results. None of these suggestions are new ones since they have been recommended by JIU since 1982.

137. It is also recognized that success in national evaluation capacity development requires a more holistic approach that starts at the design stage of programmes and continues through implementation, monitoring and then evaluation. In that regard, while the evaluation functions of United Nations organizations do have a key role to play in strengthening capacity, equally important is the role of programme units in evaluation capacity development. Evaluation is part of the governance and public-sector management of countries and hence should be made an important component of support in these areas.[67]

138. There have been several United Nations resolutions calling for national evaluation capacity development, dating back to 1989.[68] There have been efforts made, but these are scattered and the results thereof are not known. What is needed at this stage is for the organizations of the United Nations system to initiate a range of well-thought-out interventions for collaboration and partnership in strengthening national capacities for evaluation and to provide the necessary framework, resources and incentive to do this. That step would need to be preceded by an appropriately designed study on the status of the overall United Nations system effort to strengthen national evaluation capacities. This could be undertaken under the leadership of the JIU as a sequel to work done thereon in the past. Efforts to be undertaken should be tied to the 2015 year of evaluation, which focuses on national evaluation capacity.

139. The implementation of the recommendation below, taking into consideration lessons mentioned above, will enhance the sustainability of and collaboration with the evaluation function.

Recommendation 8

The Secretary-General, in his capacity as chair of the CEB, should initiate steps and support innovations for collaboration among United Nations system organizations and with other partners in strengthening national capacities for evaluation addressing accountability, learning, and knowledge development of both national and global value.

---

[67] See also United Nations System Co-operation in Developing Evaluation by Governments, JIU/REP/82/12; and Sukai Prom-Jackson, "Rebirth, Restoration, Reclamation, and Responsibilities of the Evaluation Function of Africa", in *Evaluation matter: How can we strengthen national evaluation systems – Perspectives from around the globe*, Quarterly Knowledge Publication of the African Development Bank, vol. 2, No. 3 (September 2013), pp. 76–81.

[68] General Assembly resolutions 44/211, 59/250 and 62/208 on triennial comprehensive policy review of operational activities for development of the United Nations system, and 67/226 on quadrennial comprehensive policy review.

III. THE DECENTRALIZED EVALUATION FUNCTION

A. Introduction

140. Paragraphs 24 and 25 in the introduction define decentralized evaluation for the purposes of this study. This is elaborated upon in annex V. Annex 27 in the supporting materials[69] also outlines the fit-for-purpose conceptual framework, which focuses on what is valued from decentralized evaluation and some of its special features as a function that is embedded in management. The study used this framework to answer the following main questions raised about the decentralized evaluation function.

1. Demand

Who are the main drivers of demand? What characterizes the growing focus and demands on the decentralized evaluation function? How well-positioned is the decentralized evaluation function to respond to any growing focus/demand?

2. Purpose, comparative value and normative framework

What is the nature of the decentralized evaluation function of the United Nations system? What is its purpose and what role does it play in the United Nations system architecture? What are its distinguishing features, its normative framework and standards? How does it add value in the United Nations evaluation system?

3. Structural arrangement

How is decentralized evaluation structured to support decision-making at: (a) the project level and (b) the institutional level as part of a defined plan for department-level operations? Where is evaluation located: headquarters, regional offices, country offices or technical departments?

4. Institutional framework and support systems

What is the nature of the institutional framework and support systems for the decentralized evaluation function: (a) evaluation policy, strategy, policy implementation; (b) support systems for policy implementation, and the management, conduct and quality assurance; (c) resources and financing; and (d) professional networks akin to UNEG?

5. Performance

What characterizes the performance of the decentralized evaluation function: (a) focus and coverage; (b) coherence and alignments; (c) mechanisms in place to enhance impartiality; (d) quality of reports and quality-assurance system and competencies; and (e) utility — conditions in place to enhance use, level of use and nature of use and effect of use?

6. Relevance and strategic positioning moving forward

What is the level of readiness or capacity for addressing emergent changes and demands?

---

[69] See www.unjiu.org/en/Pages/default.aspx.

B. Institutional and normative framework

Conclusion 14: Currently, decentralized evaluation at the United Nations system-wide level is not fully supported by well-defined institutional frameworks, systems and policies. Most if not all of the 28 JIU participating organizations carry out decentralized evaluations and these evaluations fall into two distinct types. For 11 JIU participating organizations, decentralized evaluation is influenced to varying degrees by the organization's systems, policies and procedures for evaluation. For the others, this is not typically the case.

141. Before describing the two types or models of the decentralized function, the following is a short note on the evolution of the decentralized function. In the early years, the United Nations evaluation system was characterized by a predominance of *self-evaluation* conducted by various organizations to support projects and programmes. This was driven to a large degree by a need for project-specific improvement of development cooperation activities. It is interesting to note that the definition of evaluation at that time also mirrors this goal of the function.[70] The primary objectives at that time for evaluation were: to cover the entire range of the organizations' activities; to provide a continuous process of rapid information feedback on emerging results and use it to quickly adjust objectives and strategy and improve future activities; to integrate evaluation as a normal part of the management cycle and as a participative process that could increase staff commitment and involvement and enhance the development of a learning organization.[71]

142. It was also presumed that, when developed by the managers responsible for the projects, the evaluations were more likely to be accepted and acted upon. Furthermore, when incorporated into the normal managerial processes, it reduced the "add-on cost" compared to a more elaborate and sophisticated evaluation systems.

143. The drawbacks recognized during this period were the quality of evaluations by staff with no professional background on evaluation and the issue of whether staff could objectively evaluate their own work. Thus, with the increasing demands for accountability with independence and professional quality coupled with the demand for evaluation to support corporate-level decision-making in the 2000s, the built-in self-evaluation (*now described as decentralized evaluation to distinguish it from the central evaluation*) became a less visible player in the overall architecture of the evaluation function. The focus shifted to the central evaluation function. The evidence drawn for this study shows an increasing need to balance the two functions as part of an integrative system responding to demand and to the emerging priorities of the development landscape.

---

[70] For a glossary of evaluation terms, see JIU/REP/78/5. See also JIU/REP/91/6.

[71] Second report on evaluation in the United Nations system, JIU/REP/81/6.

Models and archetypes of the decentralized evaluation function in the United Nations system

144. The analysis of the various archetypes of decentralized evaluation shows six variants. Not all are mutually exclusive and some are unclear as to who is accountable for quality when the production of the report is under the control of the central evaluation function. Organizations can be seen to be implementing any two of the archetypes at one time. Annex V (b) provides details of the six archetypes. There is a seventh one, under which evaluation is managed and conducted by donors supporting the funding of projects. This is not included in the chart. Overall, the six types can be categorized into two models described in this study as: (a) ad hoc and discretionary decentralized evaluation; and (b) decentralized, with an institutional framework having planned evaluations based on defined criteria for coverage and commissioned to external parties. These two types are described below.

145. **Ad hoc and discretionary decentralized evaluation.** Self-evaluations, carried out on a discretionary basis by staff as project managers with approval and use by their line managers, have continue to exist throughout the 27 United Nations system organizations[72] and operate as part of the programme management function. They can best be described as *ad hoc* and discretionary decentralized evaluations. They are generally not guided by strategic choices but are initiated at the discretion of staff and managers. In some organizations they are acknowledged in the evaluation policy as self-evaluation but in most cases, they are not guided by any specified institutional framework for their operation. In some cases, they are recognized as part of the Results-Based Management (RBM) system but not in a systematic fashion. Evidence on these *ad hoc* decentralized evaluations is difficult to assess, since they are not consistently reported or captured by the organization's evaluation and oversight systems. They represent a vast pool of the knowledge in the United Nations system supporting the organization "to create a culture where evaluations are seen as learning tools".[73] This evaluation and learning culture is most significant for the development and effectiveness of the United Nations organization. *The United Nations system is, however, missing out on the opportunity to galvanize and make use of this knowledge asset and to empower staff involved in such evaluations.*

146. Reported information suggests that these evaluations are increasingly carried out to meet the requirements of donors supplying extra budgetary funding for special projects and programmes. Thus, the prime focus has become support for accountability and, as such, is different from the original purpose of formative evaluation for learning and for project improvement as outlined in paragraph 141 above. Given its importance, the enhancement of this decentralized evaluation system would need to be based on: a solid understanding of its purpose, role and comparative added value for the United Nations system; standards that would enhance its formative role; systems to enhance and not "disempower" widespread staff engagement (given staff who may not have the professional technical skills

---

[72] UNOPS does not carry out evaluations.

[73] Accountability frameworks in the United Nations system, JIU/REP/2011/5.

of evaluators as defined in the UNEG Core Competencies for Evaluators of the United Nations system); appropriate training and learning programmes across the United Nations system on evaluation; and the development or compilation of methodologies for formative evaluation to respond to the need for the continuous improvement of projects.

147. In supporting widespread institutionalization including the development of an evaluation culture and harmonization of evaluation methodologies, UNEG has in the past initiated a system-wide face-to-face training programme for staff of the United Nations system as well as partners. This has been discontinued due to limited UNEG capacity. Ongoing, however, is widespread training on evaluation in the various United Nations organizations. There is a need for consolidated information on the diverse training and guidance materials on evaluation across the United Nations system and sharing of lessons on good practices. Such training by individual and system-wide bodies such as the United Nations Staff College must be supported by expert knowledge developed from experience of evaluation in the United Nations system to render the training relevant.

148. **It is most important that all decentralized evaluations are captured and reported in the organizational annual report on evaluation to highlight the knowledge assets of the United Nations and the broad staff involvement in building the learning and evaluation culture. A system should be developed for collaboration across United Nations organizations to enhance staff capacity for managing and conducting evaluations of various types**.

149. **Decentralized evaluation with an institutional framework.** Next are decentralized evaluations that are influenced by a well-defined institutional framework with defined rules and with a system of incentives and communication to support a consistent and strategic approach to the planning, implementation and reporting of such evaluations. The framework for such decentralized functions is characterized as follows, with due regard to variations in scope among organizations: (a) the existing *evaluation policy* defines the role of the decentralized evaluation function; (b) the evaluations *are managed in* various departments and country offices; (c) there is *planning* for such evaluations; (d) they *operate* under a set of standards; (e) there is *systematic quality-assurance or control and quality-enhancement mechanisms*; and (f) there is reporting on compliance with policy and with the development of the function. Such functions vary in the level of involvement of the central evaluation office, but in all cases they are operate under line management. The central function provides quality enhancement and/or quality control of varying types. This is outlined in annex V (b) on models of the decentralized evaluation function.

150. Eleven organizations (ILO, UNAIDS, UNDP, UNESCO, UNFPA, UNICEF, UNRWA, UN-Women, WHO, WFP and WIPO)[74] of the 28 JIU participating organizations have such frameworks in place, to varying degrees. Ten of these organizations (except UNAIDS) responded to a questionnaire and provided data for

---

[74] WIPO provided limited data. It is nevertheless included in the analysis.

analysis. The results of the data analysis from the 10 organizations provide the basis for the following findings, conclusions, and recommendations.

151. For these 10 organizations, decentralized evaluations are commissioned by their various programme offices, including technical or programme departments, thematic clusters and regional offices at headquarters, and take place in regional centres and in over 150 country offices or technical centres. Within some of these organizations, some recent strengthening of the decentralized evaluation functions is evident and an emerging range of initiatives by the various management units and evaluation units indicate an increasing focus on decentralized evaluation and reflections on the value it could add to organizations (see supporting materials on the JIU website).[75]

Conclusion 15: The whole decentralized evaluation system is missing a well-defined and appropriate normative framework and standards that are fit-for-purpose and would enhance effective management and decision-making.

152. While it is true that a considerable number of the UNEG norms and standards are applicable, they were not developed bearing in mind the purpose and nature of decentralized evaluation as an integral part of line management and the formative evaluation system and its comparative added value as part of the organizations' evaluation systems. The role of developing norms, appropriate valuation and standards that are fit-for-purpose for decentralized evaluation, in ways complementary to the UNEG norms and standards, could be part of work of the regional United Nations evaluation network of evaluation experts working on decentralized evaluation issues. The longest-established network, which has been in operation for eight years, is the United Nations Evaluation Development Group for Asia and the Pacific. Enhanced support for and engagement with such evaluation groups would be significant for the development of norms, standard and methodologies fit for the purposes of the decentralized function. **Executive heads should support the development of an appropriate normative framework and standards for the decentralized evaluation function and methodologies that are fit for-purpose, and should provide the appropriate incentive system to empower large-scale involvement of staff in evaluation and learning.**

C. Policies, implementation and institutionalization

Conclusion 16: The development of policies for decentralized evaluation has been important in guiding the development of the decentralized function in 10 organizations studied in this project. For these organizations, a considerable number of measures have been put in place to support policy implementation. The level of policy implementation and institutionalization varies among the organizations. Two key areas missing in policy formulation has been enhancing coherence among United Nations organizations and alignment with national systems. As a

---

[75] See www.unjiu.org/en/Pages/default.aspx.

result there is a very low level of development in these two areas in the decentralized function across organizations. As indicated in previous chapters, these two areas reflect similar lacunae at the central evaluation level.

153. In 8 out of the 10 organizations examined that support decentralized evaluation, the evaluation policy provides details on the role of, and support to, decentralized evaluation, but is limited in addressing coherence in the United Nations system and alignment with national systems and capacities. Across the 10 organizations examined, decentralized evaluations are commissioned at all levels, although the levels vary between organizations, depending upon how the organizations are structured and the degree of decentralization. Examination of the evaluation policies of these organizations shows that the policies provide comprehensive details on both the role of, and support to, decentralized evaluation. Issues covered include: (a) the importance and added value and the role and objectives of decentralized evaluation within the organization; (b) the arrangements for planning, management and financing of decentralized evaluation; and (c) some alignments and synergies considered critical for efficiency and sustainability.

154. Areas where centralized evaluation policies have not focused on are: (a) coordination and coherence of decentralized evaluation across United Nations organizations and the linkage between organization-commissioned decentralized evaluations and evaluation by UNDAF; and (b) the role of decentralized evaluation in supporting the development of national evaluation capacity. Inspectors find that it is in these areas that the existing systems are weakest; attesting to the importance of evaluation policies in defining what gets done. **Executive heads should ensure that the evaluation policies of their organizations fully address United Nations reform for coherence and modes of collaboration across organizations, and also fully articulate the role of the evaluation function in supporting national evaluation capacity development or other important priority areas for the effectiveness and impact of the function.**

155. In enhancing policy implementation, 8 out of the 10 organizations studied with decentralized systems have invested in supporting decentralized evaluation, mostly through the development of guidance and provision of technical support. Furthermore, they have undertaken a wide range of support activities, of which the following stand out: awareness-raising, organized meetings and workshops, special initiatives to develop a common institutional basis, capacity development via handbooks, quality assessments, training and the development of communities of practice or learning groups and networks. It is reported that two of the most effective forms of policy implementation is in engaging leadership of the organization and reporting on the progress of decentralized function to executive boards, thereby making the decentralized evaluation work and its issue visible for decision-making and institutional recognition.

156. Most valued in supporting policy implementation for decentralized functions has been the role of the central evaluation function — in particular those

of UNDP, UNICEF, UNFPA and UN-Women. They have developed pertinent standards, guidance and operational procedures for decentralized evaluations. The Inspectors also learned that of increasing importance is the role of the regional United Nations evaluation networks. An exemplary practice quoted is the joint training programme organized by the the United Nations Evaluation Development Group for Asia and the Pacific which has been instrumental in developing competencies important for policy implementation.

Conclusion 17: The degree to which relevant evaluation policies have been institutionalized varies and there are few central evaluation functions that show clear strategies on how to support decentralized evaluation or what kind of results could be expected in terms of the performance of the decentralized evaluation system.

157. Across the 10 organizations studied, the judgment of those leading on decentralized evaluation on the degree to which relevant policy has been institutionalized suggests varying levels of success. Based on analysis of the data from the questionnaire on decentralized evaluation, five organizations manifest high levels of institutionalization in terms of the policies and systems in place for implementation of the policy; they are ILO, UNDP, UNFPA, UNICEF and UN-Women.

158. A range of mechanisms have been put in place to support institutionalization and they include: guidelines, manuals and handbooks, an expert roster of consultants, evaluation database or resource platforms, follow-up response systems and websites for evaluation. Organizations vary in the degree to which these mechanisms have supported institutionalization. The usefulness of an existing roster of expert consultants appears to be limited but reportedly is of great need for the following reason: evaluations are commissioned on the basis of using independent expert consultants of diverse background and language groups to conduct the evaluation. Thus, the limitations in the number and types of evaluators and experts on the existing rosters supporting decentralized evaluation are a major problem for the impartiality and credibility of the decentralized function. The limited funding of the function also means that some consultants are not affordable. But it is reported that, in general, there are not many qualified evaluation experts on the consultancy market.

159. Organizations have recently introduced new initiatives to support institutionalization. UN-Women has introduced an online manager's guide for evaluation and has conducted training around the globe on gender-responsive evaluation for hundreds of evaluators. UNDP has developed web-based training for staff and managers and is focusing on developing strategic alliances with the decentralized function as well as with national systems. UNICEF has always taken a lead role in training on evaluation. It currently has two online training courses and has also supported the development of an online training course which is now rolled out through the EvalPartner's website (MYM&E). ILO has a certification programme for managers. UNDP, UNESCO, UNFPA, UNICEF, UN-Women and WHO have come up with various innovations in developing systems for quality assurance or for tracking implementation. ICAO, UNESCO and UNRWA have

started the process of compiling all existing self-evaluations and/or doing meta-analyses. One of the biggest organizational-wide initiatives started in 2013 as part of organizational reform was done by WHO with the development of the Global Network on Evaluation.

160. An interesting effort is being made by the central evaluation function of WFP in 2014, which, after years of limited association with the decentralized evaluation function, is developing a model to support decentralized evaluation through an innovative approach to management and quality assurance, with decentralized evaluation temporarily under the control of the central evaluation unit pending the development of a full decentralized evaluation enabling framework and capacity development. This planned approach, with central evaluation control, has identical elements with the models used by UN-Women, and the existing and emerging practices of FAO, UNEP and UNIDO.[76] Among these organizations, there are variations in the degree and type of quality control and in the use of the final evaluation reports at central or decentralized levels. Further analysis is needed to highlight some of the emerging patterns of decentralized evaluation tied to organizational structures, level of development and capacities, at both central and decentralized levels.

161. Organizations have also noted the key challenges to institutionalization to include: inadequate funding, inadequate senior management support, a weak culture of evaluation, and poor monitoring and data systems and evaluability of programmes. Some of these are also true of the central function, as discussed above. The section below highlights some of these findings at the decentralized level.

D. Resources

Conclusion 18: For most organizations, information on budget allocation or the normative basis of allocation is not fully defined or transparent. Expenditure is judged to be inadequate and they are very low relative to the financing of decentralized evaluation in bilateral agencies. Suggestions made for more efficient management include new ways of working across organizational boundaries and working in interconnected ways, including defining evaluation plans within the framework of UNDAF.

162. In response to questions on challenges for strengthening the decentralized evaluation function, 9 of the 10 organizations highlighted inadequate funding to be among the biggest obstacles. The evaluation policies of seven organizations identify the nature of budget allocation for decentralized evaluation. The source of financing is generally a percentage of the project or programme budget for four organizations. Information on actual amounts provided to decentralized evaluation is difficult to obtain: 4 of the 10 have provided information on specific amounts dedicated to decentralized evaluations. Others have budget mixed with other

---

[76] All three indicated that they do not have decentralized evaluation functions; evaluations are initiated by project managers and conducted in field locations but managed or quality assured by an evaluation in the central unit and the report is issued by the central unit and this is thus recognized as a central evaluation. Further clarity is needed on the various models.

related functions of monitoring and assessments. Organizations spend on average US$ 35,000 per evaluation (variable cost not including staff time) with a range of US$ 20,000 to US$ 55,000. While the amounts are generally higher than those spent on the UNDAF evaluations, which are much broader in scope (US$ 30,000 per UNDAF evaluation), they are low relative to expenditure dedicated to decentralized evaluations by bilateral donor agencies.[77] This raises questions about the quality of the evaluations.

163. While funding is reportedly inadequate and more is needed, some organizations have proposed, as a priority, new ways of managing costs, including an integrated strategy for using staff resources and for working together across organizations on systems and mechanisms for the function.

- According to UNDP, there is room to *enhance efficiency and combine resources, especially at country level or when addressing common themes; create greater results and evaluation planning and coherence; and build expert capacity together*. UNDAF provides an opportunity for doing this when appropriately conceptualized and planned within the framework of the One Programme and the principles of UNDAF.
- ICAO and UNRWA have proposed the *need to use the many guidelines produced by each organization and develop a compendium of tools for use across organizations*.
- Others have suggested the need *to review all existing quality-assurance and tracking systems to develop lessons that are common and specific to organizations and develop a common framework*.
- UNODC has suggested that decentralized and small organizations should use a common institution to carry out all quality assessment of evaluation reports, similar to the experience of the Department for International Development in the United Kingdom.
- Others have suggested strengthening and using national systems while in parallel enhancing their capacity.

164. A forum across organizations on this set of issues for coordination and collaboration across decentralized evaluation functions, involving representatives of the evaluation units and the managers and administrators of decentralized evaluation, would provide an important platform for initiating collaboration at the decentralized evaluation level. **The executive heads of United Nations organizations should enhance collaboration and efficiency in developing decentralized evaluation functions that respect the agenda of United Nations reform for coherence, limiting transaction costs and enhancing national ownership.**

---

[77] See United Kingdom, Department for International Development and UK Aid, "Rapid Review of Embedding Evaluation in the UK Department for International Development: Final Report" (February 2014).

E. Credibility (impartiality and quality) and utility

Impartiality

Conclusion 19: Progress has been made in enhancing the impartiality of the decentralized function. A wide range of systems and mechanisms have been put in place in 8 of the 10 organizations. These organizations use a combination of nine mechanisms designed to limit bias and control by management and to enhance transparency and trust to major stakeholders. Doing so is important to meet the expectations ensuing from the accountability demands on the function.

165. The credibility of the decentralized evaluation has been a topic of focus given the fact that evaluations are planned, managed and conducted within the management unit and approval and reporting is to or through the manager and thus arguably not independent. Considering the nature of the structure of decentralized evaluation it does not naturally meet the criteria of structural or functional independence. However, to limit management bias in the evidence, organizations have put in place a wide range of measures and mechanisms to enhance impartiality, objectivity, and transparency in decentralized evaluations. The priority is placed on the following: (a) use of external consultants; (b) stakeholder involvement; (c) use of evaluation management groups or evaluation committees; and (d) application of ethical codes of conduct and guidelines on behavioural independence. These mechanisms have played a key role in enhancing credibility, transparency, and ownership.

166. For many organizations, external consultants are used in decentralized evaluation to indicate that the evaluation is independent. In a large number of known cases, this mechanism does not ensure independence, given that the final report is approved by the manager(s) who commissioned the report. In many cases, some bilateral organizations have sought to carry out the evaluations of their funded project using their own independent means. Some organizations now have the central office doing the quality enhancement and control and ensuring that the evaluation meets criteria of quality with independence. This is the model applied by UN-Women. There is no doubt that it is the focus on accountability that is driving the wide range of efforts at the decentralized level seeking to be independent. Ensuring that the decentralized evaluation is fit for purpose must begin with reconsidering the importance of this function in first and foremost supporting the conduct of formative evaluation for improvement and for making necessary corrections for the eventual attainment of results. The United Nations system needs to consider such in developing its strategy and prioritization of the role of the evaluation function.

Quality

Conclusion 20: Eight out of the 10 JIU participating organizations have invested in evaluation skills capacity enhancement to support the quality of decentralized evaluations. More is, however, needed for enhancing the transparency and quality of evaluation reports using more objective and systematic methods that address, to a reasonable degree, both technical standards and effective management.

167. A key area of focus in enhancing the quality of decentralized evaluation has been in employing specialists in evaluation as staff to enhance quality. For the 10 organizations reviewed, monitoring and evaluation specialists are found at headquarters and at the regional and country levels. Their numbers have been increasing in recent years (UNDP, UNICEF, UNFPA, UN-Women and WFP), although arguably they are still not sufficient. In many cases, however, these specialists spend less than 25 per cent of their time on evaluation. Most of their time is instead devoted to other project cycle management tasks: planning, monitoring, reviews and other assessments. While it is reasonable for such specialists to support systems development and enhance the RBM system and evaluability both for the decentralized and the central function, excessive use in these areas overshadows their effective use as evaluation specialists in enhancing the decentralized evaluation function and its development. In the light of emerging changes and demands on the decentralized function, there is need for a reconsideration of the number and the effective use of monitoring and evaluation specialists to support core decentralized evaluation.

168. UN-Women, on the other hand, has a different model where the regional evaluation specialists supporting decentralized evaluation report both to the director of the evaluation office and the director of the regional office under a shared accountability framework and have a key role to enhance the quality of decentralized evaluation. This model has the advantage of focusing the specialist on evaluation. The effectiveness of this system of joint accountability is yet to be tested in UN-Women.

Conclusion 21: On average, based on information from 10 organizations studied, the majority of decentralized evaluations of these organizations has improved[78] but is still below average quality and more needs to be done in system development to enhance the development of solid information about the quality of decentralized evaluation on a system-wide basis.

169. Central evaluation units carry out a wide range of activities that are intended to support better decentralized evaluation. Overall, the most valued support for these activities is the assessments made of the quality of evaluation reports, the feedback provided on what to improve and the reporting of quality at the corporate level as part of annual reports on evaluation.

170. The 10 organizations reviewed produce about 640 evaluation reports per annum, covering output, outcome, impact, complex and other evaluations. Seven of the 10 JIU participating organizations have carried out in-depth and systematic assessments of the quality of the decentralized evaluations reports within the past three years.[79] The systems in place have a set of common criteria tied to the UNEG quality criteria for evaluation reports, but vary in many ways, including the number and types of indicators of quality, the process of assessment and the rating schemas used.

---

[78] The analysis of improvement is based on annual reports and selected interviews over the past four years.

[79] ILO, UN-Women, UNDP, UNESCO, UNFPA, UNICEF and WHO.

171. Nevertheless, for the purposes served by the systems for their own organizations, the analysis of reported data from six organizations shows that overall between 50 and 74 per cent of the evaluations assessed were high quality.[80] This contrasts with the opinion of interviewees from the other three organizations without assessment systems, in which it is believed that the quality of decentralized evaluations was generally below average or poor. Thus, a stable and comprehensive basis for making judgments about the quality of decentralized evaluation of the United Nations system is yet to evolve across all organizations. Existing systems developed by UNDP, UNICEF and UNFPA provide exemplary bases upon which to draw for development. Other alternative methods mentioned above include the use of one central source or institution to support a number of organizations to carry out the assessment.

Use and accountability

Conclusion 22: Like the central function, the decentralized function is highly responsive to demands for accountability for results and for using evaluative evidence to improve projects and programmes. It is yet to fully balance the demands for accountability with the need to develop the evaluation and learning culture and to enhance the sharing of evaluations within United Nations agencies and among agencies at country, regional, and global levels.

172. This conclusion is a replica of that expressed for the central level. The 10 JIU participating organizations have developed a wide range of systems to enhance use, which can be grouped into six broad categories, dealing with: (a) increasing staff and managerial engagement; (b) enhancing inclusion in the process and incentives for use; (c) the involvement of partners; (d) follow-up tracking systems; (e) websites; and (f) feedback mechanisms.

173. The systems believed to be most useful in enhancing use are: having a website through which all evaluation reports may be accessed; action plans for implementation of evaluation recommendations; systems for tracking follow-up and management response covering implementation of recommendations; inclusion of evidence from decentralized evaluations in annual reports of evaluation that are presented to senior management and/or the governing body to enhance visibility; involvement of country-level stakeholders in the planning, conduct and/or management of the evaluation to enhance inclusion and ownership; and ensuring meaningful consultations and involvement of staff and management in the evaluation while safeguarding objectivity.

174. Five of the 10 organizations have established systems for tracking implementation of recommendations and this is deemed key to enhancing use in these organizations. These systems reveal differing levels of implementation of recommendations across the five organizations. Varied effects of use are reported upon, the most significant being: enhanced ability to account for use of resources

---

[80] **ILO, UNDP, UNESCO, UNICEF, UN-Women, and WHO — evidence from annual reports of the past four years show an improvement in quality over time.**

and results; quality in decision for programmes; improved programmes and projects; enhanced transparency with effect on trust; enhanced donor support; and facility in resource mobilization.

175. Similar to the findings for the central evaluation function, the reported information for the decentralized function is consistent in showing that the effect of use for the majority of organizations is higher for accountability and for programme- and project-specific improvement and much lower for the development of a learning and evaluation culture. The analysis shows a very low effect in terms of sharing of evaluations among United Nations agencies at country level, at regional level or globally across regions. There is greater demand for accountability followed by a focus on improving programmes and projects being implemented. There is less on demand for empowering staff as knowledge workers. The major users of decentralized evaluations in decreasing order of importance are donor governments followed by project managers and implementers.

F. Readiness for change and challenges: evolving demands, capacity and readiness

Conclusion 23: The existing systems put in place for decentralized evaluation in the 10 JIU participating organizations reviewed provide an opportunity for facilitating a response to the demands from global changes and the emerging imperatives of the post-2015 development agenda in particular. The system can play a significant role when decentralized evaluations begins to address challenges regarding methodologies, priorities and coherence, on the basis of a well-defined and coherent vision and strategy to avoid ad hoc approaches and interventions.

An effective response has to be strategic and coherent and based on added value

176. The Inspectors noted an increasing focus on the importance of decentralized evaluation in most organizations. This is driven by several factors. First is the growing acknowledgement that general inadequacies in the quality of decentralized evaluations limit their effective value as major sources of evidence or building blocks for central corporate evaluation. Associated with the poor quality of evaluation is the poor quality of data and information management systems to support evaluation. This calls for increased focus of the decentralized evaluation system to strengthen capacities for data-gathering, analysis and monitoring.

177. Second is the absence of best practice models and the fast global pace for development which calls for: (a) the need for continuous, formative and process evaluation focused on testing hypotheses of what works, why and how; (b) the importance of wholesale involvement of staff in evaluation to enhance the broad strata of evaluative thinking in the organization and the development of the learning organization; (c) the need for coherence and alignment of evaluation function among United Nations agencies on the ground and at all levels of the decentralized function; and (d) recognition that decentralized evaluation need to have sufficient levels of quality and credibility to limit errors in decision-making

from their evaluations and enhance confidence from the various stakeholders of the decentralized function: donors, managers, governments, project managers and implementers.

178. Some organizations have launched in 2012 and 2013 several initiatives to respond to these emerging demands. Some have launched surveys (ICAO and WFP) or conducted meta-studies (ILO and UNESCO) to learn more about the current status and results of decentralized evaluation within their organizations. WHO has one of the most comprehensive strategies for the development of the decentralized evaluation throughout its organization tied to its ongoing programme of reform on organizational decentralization.

179. An observation made is the critical need for organizations to resist using *ad hoc* tactics and to operate from a well-defined strategic framework in addressing the development of the decentralized evaluation function. Thus, a decision to increase the load of the decentralized function must take into account existing capacity. Establishing a priority for developing the learning organization and culture must be done with a view to establishing the appropriate incentive system. To avoid the risk of an incoherent and imbalanced development, all factors affecting the development of the decentralized function need to be assessed and a coherent plan developed for moving forward to limit risk.

The level of readiness of the evaluation function for changes and challenges facing the United Nations system is rudimentary

180. The Inspectors did an assessment of the readiness of the decentralized function based on a set of significant criteria. The decentralized evaluation manifests the following strengths and weaknesses.

Major strengths

181. There is extensive project evaluation by staff organization-wide with variations between organizations. This provides a base for developing the learning organization when the effort is galvanized and enhanced using an appropriate methodology and incentive system that values risk taking and innovation. There is average alignment of decentralized evaluation with other analytic systems (monitoring, reviews, assessments and process evaluations) to allow rapidity in providing evidence for ongoing decision-making processes at programme level. There is a growing pool of high-calibre evaluation specialists and advisors involved in the field of decentralized evaluation and this needs to be fully expanded and supported.

Major challenges

182. There is an absence of a clearly defined vision and strategy to address the emerging trends outlined above and, thus, when activities are initiated they tend to be ad hoc and implications for the system as a whole and risk factors are generally not fully considered.

183. The methodologies for fast-paced learning in emergent and transformative systems would require the reintroduction of more formative evaluations and other methods, such as real-time evaluation. Likewise, the nature of systems operation and the non-structured nature of the development would have implications for the adoption of non-traditional evaluation models and methods. The absence of extensive statistical databases and sound information systems is critical. This would require strengthening United Nations organization and national statistical capacities.

184. There is an increasing demand for coherence among United Nations organizations. There is limited joint evaluation, even in the context of UNDAF which, as noted, has been in operation for 10 years now. The set of institutional constraints identified for joint evaluation due to the separate governance structures of the organizations would have to be fully addressed and managed. This would require cutting through bureaucratic boundaries and enhancing interconnections using highly proactive ways of doing things together. The determination to do so is captured in the words of one interviewee: *where there is a will there is a way.*

185. There is less consolidation of all evaluations at country level; the record shows a large pool of such evaluations that, when compiled, would provide an extensive knowledge base to guide development in countries. As an illustration, evaluations at country level for a minimal count of 150 countries with United Nations operations include: the decentralized evaluations and mid-term reviews conducted by each of the United Nations organizations on their own programmes as well as by their regional offices; UNDAF evaluations under the management of the Resident Coordinator and the United Nations country team; country-level evaluations conducted by central evaluation offices; evaluations conducted by various partners on associated projects; and other evaluations such as those done by the Global Environment Facility (GEF) and other development partners. Thus, the existing knowledge pool for any one country, with varying levels of quality, is large but not centrally accessible to the United Nations and far from being accessible to governments and interested partners. The proposed 2015 independent system-wide meta-evaluation and meta-synthesis of the UNDAF represents a much needed effort in maximizing the use of evaluations at the decentralized level.

186. There is minimal alignment within this pool of evaluations at country level. The Inspectors were told of one case of alignment between a country-level evaluation of Mongolia by UNDP and the UNDAF evaluation, which was judged to be a difficult experience in alignment. Interviews conducted suggest that successful alignment would be dependent, to a large degree on national governments taking a leadership role in ensuring that the United Nations system collaborates and consolidates its work. Ffforts directed at supporting national capacity development would have to consider this very important point including the role of government and national partners in evaluating the longer term impact of the work in the United Nations system.

187. The less-than-adequate learning organizations and culture of evaluation and information-sharing is a challenge to be addressed by senior management, which

needs to be taking a more proactive leadership role in evaluation, as opposed to a reactive role.

188. Limited alignment with national systems is an important area to be addressed and it could build on interconnections among organizations with the national government and institutions taking on a leadership role. Most governments are lacking an external independent office. Support in the development of such independent systems in countries would greatly enhance credibility in using national systems for evaluation.

189. The implementation of the recommendation below will enhance the coherence and integrity of the decentralized function.

Recommendation 9

The executive heads of United Nations system organizations should develop the institutional framework and necessary support systems to enhance the quality and added value of decentralized evaluation and the role it could play in supporting the United Nations system to address emerging challenges, including those of the post-2015 development agenda, and to enhance coherence and alignments in evaluation within and across United Nations system organizations, and with national institutions.

Annex I: Results framework of the evaluation function in the United Nations system

(Numbers listed refer to items in the Maturity Matrix of the Central Evaluation Function)

** Areas not assessed in the 2014 study. It, however, assessed readiness for impact assessment and overall value of the function*

RESULTS HIERARCHY	INDICATORS
Impact on the United Nations system*	UN system goals achieved: development, humanitarian, normative, peace and security
Intermediate Impact: *Impact of use of evaluation products**	• Enhanced credibility of the UN system • Effective oversight and governance by governing bodies and by key stakeholders – national and development partners • Enhanced quality of strategic plans, policies, programmes, projects and enhanced institutional reform • Enhanced leadership role of UN system on global landscape (knowledge, decision- making, and governance)
Outcome: *High level use of evaluation*	• Enhanced accountability for results and resources • Enhanced learning and improvement: High level use of recommendations of evaluation and lessons learned (79) to Inform decision-making and improvement of strategies, programmes and activities *(81 82 83 84)* • Enhanced contribution to advancing development evaluation and to global knowledge and impact (43) • Increased collaboration, efficiency and alignments in UN system evaluation
Output: *High quality outputs and institutional capacity*	• Coverage: Comprehensive, relevant, and adequate to support strategic decision-making, and development of policies, programmes, projects *(item 30)(core)* • High quality and credible evaluation reports (impartial, valid, reliable, inclusive perspectives) (72) (core) • High quality methodologies to advance development evaluation (65) • Enhanced Knowledge management: mining evaluations, synthesis, lessons learned and sharing internally and externally (77, 78) • Strong learning organization and evaluation capacities: Enhanced capacity and culture for results, learning and improvement and knowledge development and use. *(23,21,26,34,41,43,54,65,66,76,77,78,51,79,81,82,83,84* • Readiness and engagement for global challenges *(37,9,22,7,24,81,86,28,41,43,65,77,78,34,36,37,*

Inputs and activities: High quality systems, processes, procedures, tool, mechanisms

[redacted]

I. Relevance, Responsiveness, Adaptability, Readiness for change and challenges

- Planning, management and conduct of evaluations
- Reporting to governing bodies or management
- Lessons learned and synthesis
- Responding to reforms, mandates and challenges
  -Responding to UN Reform, Global Challenges, national capacity dev't human/gender equality and advance development evaluation
- Renewal and advancement
  -Self-assessment
  -External assessment and peer review
  -Policy review and update

II. Independence, Impartiality, Inclusivity

- Institutional mechanisms
- Functional independence
- Professional technical independence
- Behavioral independence

III. Quality- Technical and Managerial

- Staff competencies
- Standards for evaluation report
- Reference group
- External advisory panels
- Staff training and professional development

IV. Utility

- Conditions and measures to enhance use
  -Management response and follow-up, and tracking system
- Dissemination, communication and knowledge management
  -Development of analytical works and publications

[redacted]

Foundations
Charter
Resolutions
Mandate
Accountability
Framework

Institutional system
Policy
Norms & standards (UNEG et al)
Guidelines (UNEG)
RBM and accountability system
KM & learning system
Guidelines and handbooks by evaluation units

Resources
Regular budget
Extra-budget
Human Capital
Social Capital

Communities of Practice
UNEG
Regional UN evaluation groups

Organizational Culture for Accountability and Learning
Leadership
Incentives
Networks

Governance Arrangements- Roles & Responsibilities
Legislative
Executive
Evaluation

Evaluation Architecture
Alignments with systems for oversight/research/Knowledge Management

Annex II: Components and sub-components of the evaluation function in the United Nations system

A: Demand for Evaluation

Demand and Intention of demand

Organizational context

- Nature and level of demand (opened-ended questions)(see also under components III and VI)

B: Supply, Adaptation, Growth

COMPONENT: Enhancing the Enabling Environment – Organizational & Institutional Framework and Support for Evaluation and its Adequacy in Responding to Demand

- **Evaluation architecture:** (*how organized for the management and conduct of evaluation and alignment with other oversight, knowledge development and learning units)*
- **Governance:** (*Roles of Legislative, Executive, and Evaluation unit)*
- **Mandate, Vision, Strategy, and Policy for evaluation**
- **Supporting implementation of the evaluation policy:** *(promulgation, guidance, policy monitoring, feedback mechanisms)*
- **Resources (norms and adequacy), and financing**
- **Organizational culture for results, accountability, evaluation and learning:** (*RBM framework, leadership, and development of learning organization-i.e. support employee continuous learning; use and action from evaluation, reflection, use of critical inquiry and divergent thinking in operation, training for understanding at all levels, incentive system for risk taking with ideas to be tested; access to information, knowledge sharing and building communities of practice*)

COMPONENT: Enhancing Relevance (coverage) , Efficiency and Adaptability And Readiness for Change and Challenges and Moving Forward into the Future

- **Relevance** in covering programmes and projects of the organization (evaluation planning and coverage)
- **Adaptability and continuous improvement and growth**
- **Readiness for support:** (i)decentralized evaluation; (ii) UN Reform; (iii) national evaluation capacity development; (iv) global challenges; and (v) gender and human rights & diversity
- **Direction setting and reflections on the evaluation function moving into the future**

COMPONENT: Enhancing Credibility: Independence, Impartiality, Inclusion and Stakeholder Involvement

- **Structural independence** and policies for head of evaluation unit
- **Functional independence** – Planning, management, and delivery of outputs
- **Built in mechanisms for impartiality**: Controls and stakeholder involvement for balanced perspectives and impartiality
- **Professional/technical independence and integrity**
- **Behavioural independence**

COMPONENT: Enhancing Credibility: Validity and Reliability (Technical Quality)

- **Evaluators and evaluation teams:** Staff and consultant quality
- **Conditions in place to ensure quality and consistency** in the application of standards
- **Quality of evaluation reports** (output)

C: Results

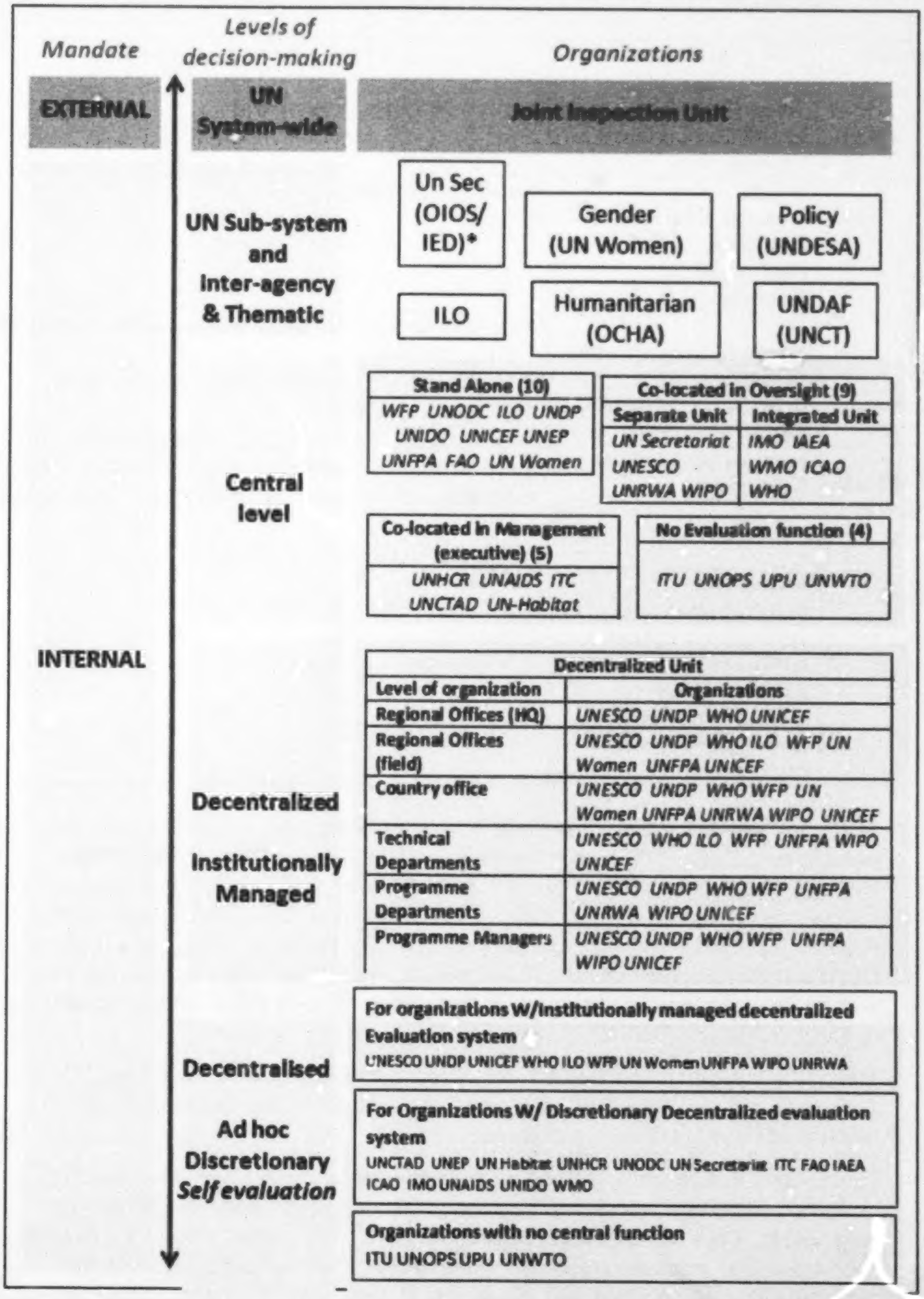
COMPONENT: Enhancing Utility and Potential Impact

1. **Conditions in place to enhance use**
2. **Outcome Level: Nature and level of use (outcome)**
3. **Impact level: Effect of use (impact)**

Annex III: JIU participating organizations from funds and programmes, specialized agencies, the United Nations Secretariat and other entities

		Mandates and Predominant (P) Mandate		
		Development	Humanitarian	Normative
Type	JIU Participating Organization			
Funds and Programmes	United Nations Conference on Trade and Development (UNCTAD)	X		X (P)
	International Trade Centre (ITC)	X (P)		
	United nations Development Programme (UNDP)	X (P)		
	United Nations Environment Programme (UNEP)	X	X	X (P)
	United Nations Population Fund (UNFPA)	X (P)	X	X
	United Nations Human Settlements Programme (Un-Habitat)	X (P)	X	X (P)
	United Nations High Commissioner for Refugees (UNHCR)	X	X (P)	X
	Children's Rights & Emergency Relief Organization (UNICEF)	X(P)	X	
	United Nations Office on Drugs and Crime (UNODC)	X (P)		X
	United Nations Relief and Works Agency for Palestine Refugees in the Near East (UNRWA)	X	X (P)	
	UN-Women	X (P)	X	X
	World Food Programme (WFP)	X(P)	X (P)	
Specialized Agencies	International Labour Organization (ILO)	X	X	X (P)
	Food and Agriculture Organization of the United Nations (FAO)	X	X	X (P)
	United Nations Organization for Education, Science and Culture (UNESCO)	X	X	X (P)
	World Health Organization (WHO)	X	X	X (P)
	International Civil Aviation Organization (ICAO)	X	X	X (P)
	International Maritime Organization (IMO)			X (P)
	International Telecommunication Union (ITU)	TC		X (P)
	Universal Postal Union (UPU)	TC		X (P)
	World Meteorological Organization (WMO)			X(P)
	World Intellectual Property Organization (WIPO)	X		X (P)
	United Nations Industrial Development Organization (UNIDO)	X (P)		X
	United Nations World Tourism Organization (UNWTO)	X (P)		X
UN Secretariat and other entities	United Nations Office of Internal Oversight Services (OIOS)	X	X (P)	X
	United Nations Office for Project Services (UNOPS)	Exec Agency		
	International Atomic Energy Agency (IAEA)	TC		X (P)
	Joint United Nations Programme on HIV and AIDS (UNAIDS)	X	X	X

Annex IV (a): Architecture of the evaluation function in the United Nations system



Mandate

Levels of decision-making

Organizations

EXTERNAL

UN System-wide

Joint Inspection Unit

UN Sub-system and Inter-agency & Thematic

Un Sec (OIOS/ IED)*

Gender (UN Women)

Policy (UNDESA)

ILO

Humanitarian (OCHA)

UNDAF (UNCT)

INTERNAL

Central level

Stand Alone (10)
WFP UNODC ILO UNDP UNIDO UNICEF UNEP UNFPA FAO UN Women

Co-located in Oversight (9)	
Separate Unit	**Integrated Unit**
UN Secretariat	*IMO IAEA*
UNESCO	*WMO ICAO*
UNRWA WIPO	*WHO*

Co-located in Management (executive) (5)
UNHCR UNAIDS ITC UNCTAD UN-Habitat

No Evaluation function (4)
ITU UNOPS UPU UNWTO

Decentralized

Institutionally Managed

Decentralized Unit	
Level of organization	**Organizations**
Regional Offices (HQ)	*UNESCO UNDP WHO UNICEF*
Regional Offices (field)	*UNESCO UNDP WHO ILO WFP UN Women UNFPA UNICEF*
Country office	*UNESCO UNDP WHO WFP UN Women UNFPA UNRWA WIPO UNICEF*
Technical Departments	*UNESCO WHO ILO WFP UNFPA WIPO UNICEF*
Programme Departments	*UNESCO UNDP WHO WFP UNFPA UNRWA WIPO UNICEF*
Programme Managers	*UNESCO UNDP WHO WFP UNFPA WIPO UNICEF*

Decentralised

Ad hoc

Discretionary

Self evaluation

For organizations W/Institutionally managed decentralized Evaluation system

UNESCO UNDP UNICEF WHO ILO WFP UN Women UNFPA WIPO UNRWA

For Organizations W/ Discretionary Decentralized evaluation system

UNCTAD UNEP UN Habitat UNHCR UNODC UN Secretariat ITC FAO IAEA ICAO IMO UNAIDS UNIDO WMO

Organizations with no central function

ITU UNOPS UPU UNWTO

Annex IV (b): Architecture of the evaluation function of the United Nations system (details on structural arrangement for the United Nations Secretariat)

Mandate	Organizations and entities	Levels of decision-making
EXTERNAL	**Joint Inspection Unit of the United Nations System**	**United Nations System-wide**
INTERNAL	**OIOS/ Central Evaluation Function (Independent Evaluation Division))**	**Sub-System wide**
	Embedded evaluation functions of departments, offices and programmes in the United Nations Secretariat	▪ **Peace and Security departments: DPKO, DFS, DPA, ODA, OOSA** ▪ **Regional Commissions: ECA, ESCAP, ECE, ECLAC, ESCWA** ▪ **UN Offices away from HQ: UNOG, UNOV, UNON** ▪ **Humanitarian and human rights: OCHA, OHCHR** ▪ **Management support: EOSG, DGACM, DM, DSS, DPI, OLA** ▪ **Others: OHRLLS and OSAA.**
	Funds and programmes and specialized agencies with own central and decentralized evaluation functions under the oversight of OIOS	**UN-Women UNEP** **UNCTAD ILO** **UNODC ITC** **UN Habitat UNRWA**

The study did not include the United Nations Secretariat departments/offices/ programmes: Peace and Security departments (DPKO, DFS, DPA, ODA, OOSA): Regional Commissions (ECA, ESCAP, ECE, ECLAC, ESCWA): United Nations offices away from HQ (UNOG, UNOV, UNON); humanitarian and human rights (OCHA, OHCHR); Management support (EOSG, DGACM, DM, DSS, DPI, OLA) and OHRLLS and OSAA. Thus it missed covering the peace and security as well as the inter-agency coordination mechanism of humanitarian and human rights of OCHA and OHCHR.

The central evaluation function for these departments is performed by OIOS/IED. The study included OIOS/IED as part of the analysis of the central evaluation function of United Nations organizations.

The various departments, offices and programmes are described as embedded evaluation functions where the function is internal to departments, offices or programmes. They manifest an evaluation function with a mixed structure. There exist evaluation functions embedded within the programme and doing their own evaluation that are planned and conducted at their own discretion. There are separate stand-alone evaluation units manned by professional evaluators that are dedicated to evaluation only (e.g. DPI, DPKO) and responding to management demands for evaluation to guide the operations of the department. There are also standalone evaluation units dedicated to both evaluation and other functions (e.g. DGACM). There are other units with no dedicated function or capacity (e.g. DESA, DM, and DPA). Interviews and consultations were conducted with these units but the information was used to guide the scope of the study.

They departments were included in the 2010 OIOS biennial study on the evaluation capacity of the United Nations Secretariat. The JIU sought to avoid duplication of effort as well as avoid another data collection process that was raised as a concern in the initial scoping period. JIU takes note of the interest raised in an analysis of the overall purpose, structure, efficiency and fit for purpose of the evaluation function of these United Nations Secretariat departments, offices and programmes along the lines of this particular study. It welcomes an opportunity for collaboration with OIOS/IED on this.

There are eight organizations that are also part of the United Nations Secretariat this arising from the fact that they get a part of their resources from the Office of the SG. They are UN-Women, UNEP, UNCTAD, ILO, UNODC, ITC, UNRWA, UN-Habitat. These are also under the oversight of OIOS. They were included in the study in their capacity as part of the funds and programmes or specialized agencies. They have a different structural arrangement for the evaluation function typical of funds and programmes and specialized agencies and this includes a clearly defined central evaluation function and a decentralized function.

Annex V (a): Definition of selected terms of the United Nations system evaluation function: central and decentralised evaluation

The evaluation function of the United Nations system has evolved. With this evolution, there has developed various concepts and constructs and associated with this a terminology with a lot of confusion. Terms used mean different things or they have different nuances in meaning. This was one of the challenges of the study. The following defines the key concepts and constructs that were defined and used in the study relative to the central and decentralized evaluation including the notion of self-evaluation. It supports the standardization of terms used. The evaluation community is welcomed to add to this list in ways that will support the evolution of the function and its emergent features.

Central evaluation function

The central evaluation function in the United Nations system is supported by evaluation offices which operate independent of the control of *line management and programmatic unit* in the planning, management, conduct and reporting of evaluation.

In some organizations the central evaluation office operates as a standalone office and there are two types: (i) operating outside the office of the executive head; and (ii) operating as stand-alone unit in the office of the executive head. In other cases it operates in the office of the executive head as in integral part of research, policy, and strategic planning (and not stand-alone). In other organizations, it operates with other oversight offices within the office of oversight of the organization and here again there are two types: (i) operating as a separate unit in the oversight office; and (ii) operating as an integrated part of the other oversight functions.

The distinguishing feature of the central function is to produce evaluations to support broad central level organizational policy, programmatic and strategic decision-making of the organization although there are many central functions that conduct project evaluations in varying degrees.

Reporting is principally to inter-governmental bodies, high level advisory bodies, and to senior management for organization-wide changes and improvements in policies, broad programmes (country, regional, global,) and cross-cutting themes and institutional reforms.

The central functions manifest varying levels of independence on the total criteria of structural, functional, inclusion, behavioural, and professional technical independence. Levels of independence vary from low to average to high levels. This independence is not automatic but involves putting in place mechanism for structural, functional, technical, transparency and behavioural independence.

Some organizations describe the central function as *corporate* implying evaluations having a wider corporate effect. The word corporate, however, also applies to the decentralized evaluations.

Decentralized evaluation function

The decentralized evaluation is planned, managed and conducted outside the central evaluation unit. Historically, it was originally designated as *self-evaluation* but this term is dropped due to its existence even in some central functions (see term self-evaluation below).

Decentralized evaluations are embedded within programme units with line management functions responsible for the planning and implementation of projects and programmes. This includes projects and programmes at the technical, regional and country office levels or departments.

The decentralized evaluation function thus generally operates as an integral part of programme management addressing evaluation, monitoring and other related analytical activities to support decision-making associated with the various phases of the implementation of projects and programmes.

Decentralized evaluations of this type are structurally not independent of control by programme managers in the planning, management, conduct, approval and reporting of the evaluation. This applies to evaluations managed and conducted by programme staff at their discretion as well as evaluations that are commissioned or contracted to external consultants for which the final approval of the report rests with the programme managers. This also includes UNDAF evaluations.

Non-independent decentralized evaluations that are conducted by staff or are commissioned to external consultants but are under the control of line management for the planning, conduct and reporting are not independent.

Independent decentralized evaluations that are planned at the decentralized level and other levels as well, are under the control of other credible independent bodies and for which the final approval and issuing of report are independent, are considered as independent. Their credibility in such circumstance, however, depends on transparency in the involvement of key stakeholders.

Self-evaluation

This terminology raises the most confusion. The JIU made the decision in this study not to use it to categorize functions. It is used principally as an adjective.

The general position is that self-evaluation is not independent given:

(i) It is not done by the central unit. It is, however, argued that some central units are not independent when under the executive head or they manifest varying levels of independence when all factors of independence identified in the JIU maturity matrix are considered.

(ii) It is not done by an external consultant. It is argued that merely using an external consultant does not guarantee independence when the report is still under the control of the programme manager. External consultants also do not always operate with an independent perspective.

Self-evaluation is undertaken by project managers for their own use to improve projects and programmes.

(iii) It is argued that this interest driving managers means mangers will operate with integrity and not tamper with the evidence given the need to know for improvement. In this regard the issue of independence is irrelevant and replaced by management integrity and trust.

Self-evaluations used strictly for programme use (as characterised in some definitions) are rare as they also support the demands for accountability for the use of resources in various contexts and particularly in a context of raising extra-budgetary support. It is argued that it is best to use the word formative evaluation to describe evaluations that fed directly into project and programme design and improvement.

The following are both past JIU and more recent OIOS definitions of self-evaluation. Both are now categorized under the decentralized function in this study.

Self-evaluation - historical definition (from JIU/REP/81/6): *"evaluations conducted to cover the entire range of the organizations' activities, ...a continuous process of rapid information feedback on emerging results and their use to quickly adjust objectives and strategy and improve future activities, ...the integration of evaluation as a normal part of the management cycle, ...as a participative process which could increase staff commitment and involvement and enhance building a culture of evaluation and to enhance the development of a learning organization"*. This was done by all organizations in the 1980s.

Self-evaluation (OIOS List of Key Oversight Terms, Version 1.0, April 2013)

Evaluation undertaken by project and programme managers primarily for their own use. Self-evaluation focuses on the sub-programme and can be applied to all sectors, including political, legal, humanitarian, economic and social affairs, public information and common services. As an integral part of the management process, the design and conduct of the self-evaluation procedure are specified at the planning and programming stages in conjunction with the design of the relevant sub-programme. Findings are applied by the programme managers to make necessary adjustments in implementation or are fed back into the planning and programming process in the form of proposed changes in the design and/or orientation of the sub-programme or project concerned. While the results of self-evaluation are not normally reported at the intergovernmental level, the conclusions drawn *from the analytical information generated for various sub-programmes and projects within a programme can be used as inputs for assessing the programme as a whole.*

Types of decentralized evaluation in the study

This study categorises the decentralized function into two main models as outlined below. Under each of the two modes are various archetypes that are not necessarily exclusive. These archetypes are provided on the next page in annex V (b).

Ad hoc, discretionary decentralized evaluation

Conducted by staff and managers at their discretionary and not aligned with defined institutional demand. They are not tied to an institutional framework for reporting and other strategic decision at programme level. They follow the definition of the JIU and OIOS although use could also be accountability purposes as evidenced by anecdotal evidence on recent trends.

Institutionally framed evaluation for decentralized evaluation

Given the largess of the ad hoc model, this model was developed to enhance a more planned and strategic value to the decentralized evaluation function.

These are decentralized evaluations that are influenced by a well-defined institutional framework with defined rules and with a system of incentives and communication to support a consistent and strategic approach to the planning, implementation and reporting of such evaluations. The framework for such decentralized functions is characterized as follows: (i) the existing *evaluation policy* defines the role of the decentralized evaluation function; (ii) the evaluations *are managed in* various departments and country offices; (iii) there is *planning* for such evaluations; (iv) they *operate* under a set of standards; (v) there is *systematic quality assurance or control and quality enhancement mechanisms*; and (vi) there is reporting on compliance with policy and with the development of the function.

Such functions vary in the level of involvement of the central evaluation office but in all cases, they are managed by line management under which they operate. The central function provides quality enhancement and/or quality control of varying types. Organizations vary in the degree to which the factors identified for institutionalization exist.

Annex V (b): Decentralized evaluation function models and archetypes in the United Nations system

The various archetypes are not mutually exhaustive. Organizations can apply more than one archetype concurrently

Model 1: Ad hoc and Discretionary- Almost all UN Organizations			**Model 2: Well defined Institutional Framework Provided - Selected UN Organizations**		
Archetype 1:	*Archetype 2:*	*Archetype 3:*	*Archetype 4:*	*Archetype 5:*	*Archetype 6*
Planning, Management is done by programme unit.	Planning, Management is done by programme unit.	Planning and management is done by the programme unit.	Planning and management is done by the programme unit (following policy).	Planning is done by programme (following policy)	Planning is done by or with line management (following policy)
Institutional framework not defined or RBM provides this	Institutional framework not defined or RBM provides this framework.	Institutional framework not defined or RBM provides this.	Institutional framework defined in varying degrees for policy, planning, management and conduct.	Decision for what to evaluate is approved by executive management.	Decision for what to evaluate is approved by executive management. Selection is strategic. Criteria are established for what gets evaluated.
Evaluation is done by staff member(s) of project or programmes	Evaluation is externally commissioned and conducted by an external consultant selected on technical criteria. Consultant is independent from programme or project being evaluated	Conduct of evaluation is by staff or an externally commissioned consultant	Conduct of evaluation is by staff or externally commissioned consultant(s)	Evaluation is conducted by external consultant. But managed by line management and programme unit	Evaluation is managed and conducted by central evaluation unit
Quality enhancement and assurance is ad hoc	Quality enhancement and assurance is ad hoc	Quality enhancement, support and guidance by the central evaluation office(either via manuals and guidance materials)	Intensive quality enhancement, support and guidance by the central evaluation office (either via direct help desk support and/or via manuals and guidance materials and training for both evaluation and quality in project design (ie. evaluability assessments)	Quality enhancement, support and guidance by the central evaluation office mostly limited to guidance materials.	Quality enhancement, support and guidance by the central evaluation office
Final approval and clearance by line manager	Final approval and clearance of report by line management.	Final approval and issuance of report is by line management.	Approval, clearance and issuance by the central evaluation office following validation or verification or other quality control mechanism of each report	Final approval and issuance of report is by line management Assessment of quality of report for corporate reporting and compliance is done by the central evaluation unit to support policy making	Final approval and issuance of report is by the central evaluation office
Report used for own use -improvement and correction.	Report used for own use - improvement and correction.	Report used for own use - improvement and correction.	Reports for broader use	Reports for broader use	Reports for broader use and as part of pool of reports for central function.
Reports not captured in knowledge management system	Reports not captured in knowledge management system	Reports not captured in knowledge management system	Reports are captured in knowledge management system or evaluation resource platform	Reports are captured in knowledge management system or evaluation resource platform	Reports are captured in knowledge management system or evaluation resource platform

Annex VI: Business models used by 23 United Nations system organizations

Types of Models	Number of organizations with Model	Number of evaluations completed under model in 2012-2013
Model 1: Evaluation is commissioned to an outside expert or company. This external source has full accountability for quality. It reports on the evaluation. Evaluation unit's responsibility is to support the conduct of the evaluation. Quality assurance of the evaluation could be by an external evaluation management group or external Readers	2	5
Model 2: Evaluation unit is fully accountable for the evaluation. It leads the overall management, quality assurance, and completion of the report. It ensures that evaluation is meaningful to the organization.(Institutional validity) Team leadership for the conduct of the evaluation is contracted to a subject matter expert who also has an evaluation or extensive analytic background and strategic outlook. Likewise, team of experts supporting the evaluation is also contracted. Evaluation unit could also participate in the pilot or the conduct of the evaluation and in some cases, could serve as the co-team leader.	16	417
Model 3: Evaluation unit is fully accountable for the evaluation. It leads the management, quality assurance, and completion of the report and has responsibility for institutional value of the evaluation. Team leadership is by the Evaluation Uni. Co-team leadership with an external expert is also an option. Unit uses contracted expert support as needed.	10	72
Model 4: The evaluation is done internally and no outside sources are contracted	3	9
Model 5: Other models: include project being done at decentralized level but under quality control and final approval by the evaluation unit	3	35

Annex VII: Demands on the evaluation function beyond the conduct and presentation of evaluation reports

Internal Support to Organization including Decentralized Evaluation	Development of the Function for the Organization	Outreach (Global, Regional, and National)
RBM and support quality in project and programme design	Policy development Policy review and update (by external team) Policy promulgation	Engagement of Global Platform and Global Knowledge: Conferences and workshops Publications
Develop guides and tools on evaluation – organization wide	Annual Report on the evaluation function and on evaluations	Normative work, Coherence, Harmonization and knowledge sharing: UNEG and UN reform: UNEG AGM: UNEG EPE
Conduct staff training (webinars, face to face, distance learning) Conduct training for senior management Develop web-based training tools	Lessons learned from various evaluations (synthesis) for management action.	Global partnerships in methodology: Participation as member of Network of Networks on Impact Evaluation and as Member of EvalPartners
Decentralized evaluation: Quality enhancement & Quality control	Communication and dissemination	National evaluation capacity development
Follow-up Tracking Systems development and reporting	Methodology Guides for evaluations of central function	
Data systems: Platform for all evaluations of the organizations and knowledge management		Linkage for Independent System-wide evaluation

Annex VIII: Evolution of Evaluation Offices in JIU POs, 1977-2013

An Illustration of progression over time in level of development and staffing increase

(Analysis is based on data derived from multiple data sources. Data sources do not all report on the same organizations). Dates indicate year of establishment of the Function

CENTRAL EVALUATION FUNCTION											
	Level of Development						Number of Staff and Consultants in Evaluation Units				
		1980					2011	2013			2012-2013
Level of Maturity*	1977	20 organizations	Number of staff	1980-2005 (No data)	2006	2013	Number of staff (21 Orgs)	Number of professional staff (24 Orgs)	Number of general services staff	Total number of staff	Number of Consultants (17 Orgs)
							132	160	39.5	199.2	988.5 (for 2 years)
Level 5											
Level 4						UNDP	15	18	6	24	160
Transitioning to Level 4						ILO	7	5.5	1	6.5	85
						WFP	8	16	3	19	188
						UNWOMEN (2011)	8	12	2	14	X
						UNIDO	4	4	3	7	141
						UNICEF	7	14	3	17	X
						UN-OIOS	19	22	4	26	X
						UNFPA	5	6	1	7	46
						UNESCO	X	6.5	1	7.5	X
Level 3 (Major elements of function in place and routine operations ongoing)		FAO	8		UNDP	FAO	15	17	4	21	318
		UNFPA	5		WFP ITC	IAEA-	5	5	X	5.3	2
		WHO	1		UNIDO	UNEP	7	6	3	9	3
		WFP	10		UNICEF	UNAIDS	2	X			X
		UNDP	1		UNESCO	UNODC	4	3	2	5	X
		ITC	2		FAO	WIPO	2	3	2		8
		UNEP	2		UNODC						
					UN-Habitat						
Transition to Level 3 (Some elements of the system in place and mechanical operation)	WFP (1963)	ILO (1977)	3		ILO	ITC		2	1	3	7
	UNDP (1967)	UNESCO (1987)	2		UNWOMEN	WHO		4	0.5	4.8	X
	FAO (1968)	UNHCS (1979)	1		UNOIOS(1994)	UNCTAD		2	0.5	2.5	11
	WMO (1968)	WMO	2		UNFPA	UN-Habitat		2	2	4	1
	UNFPA (1972)	UNIDO	1		IAEA						
	UNEP (1975)	UNHCR (1980)	4		UNEP						
	ITC (1975)	UNICEF	2		UNCTAD						
	UNICEF (1976)	IAEA (1978/80)	10								
	UNIDO (1976)	UN/DESA (1980)	4								
	WHO (1976)										
	UN/DM(1974)										

Level 2 **(Being developed and ad hoc)**	UNESCO ILO					ICAO UNHCR IMO WMO UNRWA	2 5 2 2 2	2 7 1 1 2	0.2 1 X 0.25 X	2.45 8 0.5 0.75 2	1 13 3 X 2
Level 1 **(Non-existent/ orienting towards development of function)**	IAEA ITU IMO UNCTAD	IMO UNCTAD ITU UPU				UNOPS UPU UNWTO ITU					
DECENTRALIZED EVALUATION FUNCTION											
	Organizations with discretionary evaluations							**2013: Reported level of quality of support provided by the central evaluation unit to the decentralized function**			
	UNDP UNIDO ILO IMO UNESCO WHO UPU ITU	UNDP ITC WHO UNICEF UNEP UNFPA UNHCR FAO ICAO UNESCO WHO ILO UNIDO ITU W/UNDP				**Extensive support** *UNODC UNOIOS, UNFPA* *UNICEF ILO UNIDO WHO*			**Limited support** *UNCTAD UN-Habitat UNHCR* *UNRWA WFP FAO UCAO IMO* *UNESCO WIPO*		
						Good support *ITC UNDP UN-WOMEN* *UNAID*			**No support** ***UNEP IAEA WMO***		

Note: The conceptualization of the level of maturity or development is based on the JIU conceptual scheme on the Maturity Matrix. Analysis for 1877 and 1980 is based on JIU Reports focused on the function. Analysis for 2006 draws on many sources: (i) UNEG Task Force Quality Stamp: Baseline Synopsis of UNEG Members Checklist Answers – Based on Self-Assessment Of 23 UN Organizations of which 16 are JIU PO's. This was done against the UNEG Norms and Standards: (ii) Peer Reviews of individual organizations between 2006 and 2007) (UNDP, UNICEF); JIU reports on Oversight and Accountability Framework 2005/6. Analysis for UN Women (1911) is based on record of UNIFEM evaluation function and that of the other associated agencies focused on gender and women empowerment that operated under UNOIOS..

**Org = United Nations system organizations*

Annex IX: The learning organization: culture for results, accountability, double loop learning, sharing and use of evaluation

Analysis across organizations: Means score and number of organizations at various level of development on indicators describing the learning organization

	Mean Score	**Number of organizations performing at various level of the maturity matrix**			
		Level 1 1-2	***Level 2 3-4***	***Level 3 5-6***	***Level 4 7-8***
Organizational culture for results, accountability, learning (23) (Cote item- Overall summary rating in maturity matrix)	4.3	1	11	12	-
I. Strategic alignments, outreach, generativity					
Alignment with other support systems (21)	4.7	-	10	11	3
Support to decentralised evaluation (26)	4.8	1	11	3	8
Support to national evaluation capacity development (34)	3.5	8	9	5	2
TOTAL	**4.35**				
II.Development, contribution, transformation, creation					
Continuous assessment of function (41)	5.5	1	7	7	10
Contribution to advancing knowledge on evaluation (43)	4.8	4	8	4	8
Controls and stakeholder involve and inclusion of perspectives (54)	6.0	-	3	11	10
Methodologies – development and innovation (65)	5.3	-	9	11	4
Professional develop of staff (66)	5.1	-	12	5	7
TOTAL	**5.34**				
III.Sharing and communities of Practice					
Enhance accessibility and transparency (76)	6.4	2	3	3	16
Sharing internal (77)	5.4	1	5	10	8
Sharing external (78)	4.7	6	5	10	3
Provide access to information (51)	6.0	1	2	11	10
TOTAL	**5.6**				
IV.Double Loop learning – use and action					
Recommendation implementation (79)	4.8	7	1	11	5
Use for strategic decisions (81)	5.0	1	7	12	4
Corporate/summative use (82)	5.2	1	7	8	8
Formative use (83)	4.9	1	9	9	5
Use external to organization (other UN agencies, national, development partners) (84)	4.3	3	9	11	1
TOTAL	**4.84**				
Overall Mean Score	**4.8**				
Average number of organization at various levels		**2.1**	**7.3**	**8.8**	**6.2**

Note: *Numbers next to the indicators are the numbers of the rubrics in the maturity matrix*

Annex X: Leadership attributes and functions

Analysis across organizations: Means score and number of organizations at four levels of development

Leadership Functions	Mean Score	Number of organizations performing at various levels of the Maturity Matrix			
		Level 1 *1-2*	*Level 2* *3-4*	*Level 3* *5-6*	*Level 4* *7-8*
Overall Leadership of Senior Management*					
Core item (22) Predominant responsibility (core)	4.4	0	11	13	0
I. Vision, Strategy, Relationships					
Vision and strategy (9) *Joint responsibility*	5.2	1	7	12	4
Governance structure(6) *Joint responsibility*	5.8	0	7	6	11
II. Institutionalization of function					
Policy promulgation (12) *Joint responsibility*	5.1	1	4	16	3
Learning organization and culture for evaluation (23) *Predominant responsibility*	4.3	2	10	12	0
Related and support systems and alignments and incentives for strengthening evaluation (21) *Predominant responsibility*	4.7	0	10	11	3
Resources are dedicated, stable and adequate (15,19,50) Predominant responsibility	4.8	2	10	9	3
Ensure use of evaluation (79,81,82,83, 84,85,86) *Joint Responsibility*	4.5	2	8	14	0
III. Accountability and Independence					
Safeguarding independence and integrity in approach (61) *Predominant responsibility*	5.9	1	1	12	9
Ensuring access to information for evaluation(51)(independence and transparency) Key issue however is more on quality of data and efficiency in accessing data and information and systems in place. *Predominant responsibility*	5.9	1	2	11	10
Ensure independence of programme of work (49) *Predominant responsibility*	5.7	1	4	9	10
Accessibility and transparency (76) *Joint responsibility*	6.4	2	3	6	16
Average score	**5.0**				
Average number of organization at various levels		1	6.7	11.4	4.8

** Senior management fully understand the role and added value of evaluation, actively support and promote the function within the organization and develop the appropriate incentive system. The JIU principles for leadership and development of a culture of learning and accountability exist including leading the evaluation function: by values and example, by information and communication; by motivation, by guidance and discipline; by participation.*

Note: *Numbers next to the indicators are the numbers of the rubrics in the maturity matrix*

Annex XI: Relationship between level of non-core financing, safeguards for effective and independent use, and level of development of the function

% of Non-Core Resources for 13 organization *(with complete data sets)*	Mean score for level of development	Mean score for safeguards to independence, sustainability of function, effective use and limited transaction costs (indicator 18)	Mean Score Independence (structural, functional, behavioral, professional technical.) (Indicators 45 to 62)
Low level of non-core financing			
1-10%: UNDP UNESCO UNFPA	6.3	6.3	6.7
11-20%: WFP, UN Women IAEA UNCTAD	5.8	6.6	6.4
Total	6.1	6.5	6.5
Average level of non-core financing			
21-30%: UNIDO, ITC	5.7	5.5	5.4
31-40%: UNICEF, FAO,	6.1	4.5	6.4
Total	5.9	5.0	5.9
High level of non-core financing			
Above 40%: UNEP (46%) UN-HABITAT (82%)	5	4.2*	5.1

**Estimate of data for UN-Habitat*

Annex XII: Readiness for global changes and challenges including the imperatives of the post-2015 development agenda for sustainable development

	Mean Score	**Number of organizations performing at various levels of the maturity matrix**			
		Level 1 1-2	***Level 2 3-4***	***Level 3 5-6***	***Level 4 7-8***
Addressing Global Change and Challenges (Core Factor) (37)*	5.0	-	8	5	2
I.Strategic in Reflection and Outlook					
Vision and strategy (9)	5.2	1	7	12	4
Leadership (22)	4.4	-	11	13	0
Member state appreciation (7)	5.5	-	7	11	6
Planning for coverage/strategic and optimal (24)	5.8	-	6	7	11
Use strategic (81)	5.0	1	7	12	4
Effect of use – impact (86)	3.9	4	10	10	0
TOTAL	4.9				
II.Innovation, Adaptation, and Renewal					
Balancing activities (28)	5.6	-	6	9	9
Continuous assessment (41)	5.5	1	6	7	10
Contribution to advancing development evaluation (43)	4.8	4	8	4	8
Methodologies and innovations (65)	5.3	-	9	11	4
TOTAL	5.3				
III.Engagement including on Global Platforms					
Sharing internally (77)	5.4	1	5	10	8
Sharing externally (78)	4.7	5	5	10	3
TOTAL	5.0				
IV.Coherence, Inclusion, Ownership, Alignments					
Support to national evaluation capacity (34)	3.5	8	10	10	0
Gender Equality, HR and Inclusion (36)	5.2	1	9	5	9
Support to decentralised evaluation (37)	4.7	3	10	3	8
TOTAL	4.4				
Developing the Learning Organization (aggregate score – list items)	4.8	-	8	16	0
Mean Score across organizations	4.9				
Average Number of organizations at various levels		1.5	7.7	9.1	4.5

* *Fully cognizant of global trends and challenges. Seeking new approaches and partnerships(in planning, coverage, joint work, methodologies and context factors of consideration) and revising old partnerships and methods for cognitive diversity and new imperatives*

Note: *Numbers next to the indicators are the numbers of the rubrics in the maturity matrix*

Annex XIII: Overview of actions to be taken by participating organizations on the recommendations of the Joint Inspection Unit

JIU/REP/2014/6

		Intended impact	United Nations, its funds and programmes															Specialized agencies and IAEA														
			CEB	United Nations*	UNCTAD	ITC	UNDP	UNEP	UNFPA	UN-Habitat	UNHCR	UNICEF	UNODC	UNOPS	UNRWA	UN-Women	WFP	FAO	IAEA	ICAO	ILO	IMO	ITU	UNAIDS	UNESCO	UNIDO	UNWTO	UPU	WHO	WIPO	WMO
Report	For action		☐	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☐	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒
	For information		☒	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☒	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Recommendation 1		f	▨	▨	E	E	▨	▨	▨	▨	E	▨	▨	▨	E	▨	▨	▨	▨	E	▨	E	▨	E	▨	▨	▨	▨	E	▨	E
Recommendation 2		f	▨	E	E	E	E	E	E	E	E	E	E	▨	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
Recommendation 3		f,h	▨	L	L	L	L	L	L	L	L	L	L	▨	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
Recommendation 4		f	▨	L	L	L	L	L	L	L	L	L	L	▨	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
Recommendation 5		f	▨	E	▨	▨	E	E	E	E	E	E	E	▨	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
Recommendation 6		f	▨	E	E	E	E	E	E	E	E	E	E	▨	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
Recommendation 7		d	▨	E	E	E	E	E	E	E	E	E	E	▨	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
Recommendation 8		c	▨	E	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨
Recommendation 9		d,i	▨		E	E	E	E	E	E	E	E	E	▨	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E

Legend: L: Recommendation for decision by legislative organ E: Recommendation for action by executive head

▨ : Recommendation does not require action by the organization

Intended impact: **a:** enhanced transparency and accountability **b:** dissemination of good/best practices **c:** enhanced coordination and cooperation **d:** strengthened coherence and harmonization **e:** enhanced control and compliance **f:** enhanced effectiveness **g:** significant financial savings **h:** enhanced efficiency **i:** other.

* Includes only OIOS. Covers all entities listed in ST/SGB/2002/11 other than UNCTAD, UNODC, UNEP, UN-Habitat, UNHCR, UNRWA.